Part-of-speech sensitive suprasegmentals in Portuguese-related Creoles

Alain Kihm
CNRS – Université de Paris

In most Portuguese-related Creoles (PRCs) nouns and verbs are systematically distinguished by contrasting prosodic patterns, which are therefore part-of-speech (POS) sensitive. Because they proceed from Portuguese infinitives having discarded final /r/, verbs end in CV syllables in PRCs (except in Diu Indo-Portuguese). When suprasegmental dominance is realized as stress, verbs bear main stress on this final syllable, whereas CV-ending nouns are stressed on the penult. Only nouns ending in heavy (CVC) syllables show final stress. When tone is the main suprasegmental, different tone patterns distinguish nouns from verbs. PRCs thus depart from their lexifier where stress assignment is not POS-sensitive, and instantiate a rare typological feature: morphophonological devices exclusively devoted to tell POS apart. The present study attempts to account for this state of affairs. The starting point was variable final *r*-drop in sixteenth century Portuguese and Língua de Preto (LDP), the Basic Variety among African slaves in Portugal. In LDP and subsequent creoles, however, it only affected infinitives, thus leading to the stress/tone POS divide. I argue that the need to distinguish predicative from argumental items, especially pressing in unguided L2 learning situations, was the leading force behind a change that resulted in a major typological split.

Keywords: Basic Variety, infinitive, Língua de Preto, part of speech, rhotic, stress, suprasegmental, syllable structure, tone, unguided second language acquisition.

1. Introduction

The aim of the present work is to study a particular phenomenon related to suprasegmentals of prominence — i.e. stress and tone (see Hyman 1975: 203ff.) — in the Portuguese-related creoles (PRCs) of the Atlantic as well as the Asian area. I am referring to the fact that in most languages in both areas words belonging to different parts of speech (POS), nouns and verbs in particular, having the same syllabic structures or even identical segmental forms show systematically distinct prosodies such as differential stress or tone patterns uniquely related to the lexical category contrast. Hence minimal pairs such as

Guinea-Bissau-Casamance Kriyol *kasa* /ˈka•sa/ ‘house’ vs. *kasa* /ka•ˈsa/ ‘to marry’, Lung’ie *fala* /fa^{H}•la^{H}/ ‘speech’ vs. *fala* /fa^{L}•la^{L}/ ‘to speak’, Papiá Kristang *káza* /ˈka•za/ ‘house’ vs. *kazá* /ka•ˈza/ ‘to marry’, Papiamentu *peña* /ˈpe^{H}.ɲa^{L}/ ‘comb’ vs. *peña* /ˈpe^{L}•ɲa^{H}/ ‘to comb’ (Kihm 1994: 14; Maurer 2009: 26; Agostinho & Hyman 2021; Baxter 1988: 39; Kouwenberg 2004).[1, 2]

The issue is not trivial. To quote Bakker (2008: 145), “formal markers for word classes are cross-linguistically uncommon”.[3] Such a statement should not be misunderstood. What is uncommon is markers whose *sole* function is POS distinction, such as the endings *-o* and *-a* of Esperanto showing the word to be a noun and an adjective respectively (see Kalocsay & Waringhien 1985: 19). Not uncommon, in contrast, are markers that occur only in a given POS, e.g. nouns or verbs, to express properties that happen to be only appropriate for this POS. In Russian, for instance, a word ending in *-ov*, if not a person’s name, is sure to be a noun, as verbal paradigms do not include this suffix. Yet *-ov* is not a nounhood marker, but it is the exponent for the genitive plural of 2nd declension nouns ending in a nonpalatalized consonant, and genitive plural happens to be a morphosyntactic feature set only appropriate to nouns. So is the morphomic feature 2nd declension implying masculine gender — although not all masculine nouns belong to this inflectional class.

Another not uncommon state of affairs not to be confused with prosodic POS distinction, as in the PRC examples above, is illustrated by the following three Portuguese word-forms: *cântara* /ˈkɐ̃^{n}tɛrɐ/ ‘jug’, *cantara* /kɐ̃^{n}ˈtarɐ/ ‘s/he/it had sung’, *cantará* /kɐ̃^{n}tɐˈra/ ‘s/he/it will sing’. Segmentally identical (but for the varying vowel reductions), only the stress patterns tell them formally apart. The stress patterns, however, have nothing to do with POS. Proparoxytonic nouns such as *cântara* are numerous, but paroxytones such as *palito* /pɐˈlitu/ ‘stick’ are much more common, probably representing the default for nominals

[1] H = high tone, L = low tone. A general caveat : statements about prosodic patterns hold only of words in isolation immune from contextual interference. For instance Kriyol *N kasa* ‘I married’ has (probably secondary) stress on the first syllable /ŋka/ for reasons to be discussed later on.

[2] The same phenomenon is observed in Spanish-related Chabacano and (not so clearly) Palenquero : cf. Ternate and Cavite Chabacano *kása* ‘house’ vs. *kasá* ‘to marry’” (Sippola 2013a: 144; Sippola 2013b: 150). Lacking expertise in this field, I prefer to leave Spanish-related Creoles out of consideration. I nonetheless include Papiamentu. Although Spanish-lexified, it makes sense to annex it to the Portuguese domain (see Jacobs 2013; Maurer 2013c).

[3] Bakker is referring to segmental word-class markers in a number of pidgin languages, e.g. Russenorsk. In those pidgins, it is nearly always verbs that are marked as such by special endings, whereas nouns are unmarked or default. As we shall see, this is not true of the noun-verb contrast in PRCs.

ending in a vowel (Magalhães 2016: 108-109). The stress difference between *cantara* and *cantará* is related to the tense difference: past forms have stress on the final vowel of the theme, future forms on the tense ending (Magalhães 2016: 109-111). Vowel-final oxytonic nouns do exist: e.g. *avô* /ɐˈvo/ 'grandfather', *café* /kɐˈfɛ/ 'coffee', *javali* /ʒɐvɐˈli/ 'wild boar', *peru* /pəˈru/ 'turkey', etc. In Guinea-Bissau-Casamance Kriyol, in contrast, no word ending in an oral vowel has final stress unless it is a verb or penultimate stress unless it is a noun.[4]

True, this state of affairs is reminiscent of English pairs such as *abstract* /ˈabstrakt/ (noun, adjective) vs. *abstract* /əbˈstrakt/ (verb) (see Jespersen 1954: §5.7 for a full list). There is a fundamental difference, however. It is not the case that English verbs are systematically oxytonic, whereas nouns would never be, but for exceptions. Although numerous (about 200), such pairs — for which Jespersen provides a historical explanation — constitute an island within the lexicon and stress system of English, which system is not sensitive to POS in the majority of cases.

What makes POS-sensitive stress assignment in PRCs a significant issue, in contrast, is that, far from being a late development, it goes back to the very beginnings of pidginization / creolization, namely the Basic Variety of Middle Portuguese commonly called Língua de Preto, spoken by the African slaves brought to Portugal as soon as the second half of the fifteenth century and made known to us by the literature of the time.[5] Língua de Preto texts provide many examples of verb forms ending in vowels graphically marked for stress, e.g. *crupá* 'to blame' (P *culpar*) as opposed to unmarked and therefore paroxytonic *crupa* 'fault' (P *culpa*).

The origin of such contrasts is clear: of all the inflected forms of Portuguese verbs Língua de Preto speakers almost exclusively retained the infinitive, which they pronounced without sounding the /r/ but still stressing the final syllable, so that the location of stress in such forms became lexically determined. What is unclear, however, is why systematic *r*-drop only impinged on infinitive forms. Nouns ending in /r/ either were taken as such (e.g. *mar* 'sea',

[4] There are a few exceptionally oxytonic vowel-final nouns such as *mame* [maˈmɛ] 'mother', *pape* [paˈpɛ] 'father', *kafé*, etc. Not only are they much less numerous than in Portuguese, but most of them do not belong to what may be considered basic vocabulary, present in the language by way of the creolization process itself rather than by more or less recent borrowing. Nasal vowels in the final syllables of nouns such as in *kaleron* 'cauldron' (P *caldeirão*) receive stress, but *kaleron* is not vowel-final, being pronounced /kaleˈrõŋ/.

[5] On Basic Varieties see Klein & Perdue (1997). For Língua de Preto see Giese (1932), Teyssier (1959/2005), Kihm & Rougé (2013). A similar Spanish Basic Variety called Habla de Negro appeared in Southern Spain at the same time in comparable circumstances (see Santos Morillo 2020).

senhor ‘sir, lord’) or final /r/ was ‘covered’ by a paragogic vowel (e.g. *muyere* ‘woman’ < P *mulher*) — a process that also affected verbs to a lesser extent (e.g. *buscaro* ‘to look for’ < P *buscar*). Was then the need formally to distinguish POS the rationale for the phonological change? And if so, why is such a need so seldom manifested, or so it seems, in natural languages?

Trying to answer these questions is the purpose of the present study. It is organized as follows. First I will provide a reminder of stress assignment in Portuguese. Then I will present an overview of the suprasegmental systems of a sample of PRCs, including the Basic Variety just mentioned, illustrative of the change. I will then examine whether the linguistic environments (substrates-adstrates) the creoles grew in may have triggered or at least favoured this change. Having found the evidence for such an influence inconclusive, I will envisage the possibility that a change in the realization of final rhotics in Middle Portuguese may be the ultimate culprit. I will argue, however, first that there is no solid evidence for such an event, secondly that, even assuming it did occur, it could not have entailed the observed consequences were it not for the way creole languages come into existence.

2. Stress assignment in Portuguese: a reminder

In this section I will limit myself to reviewing those facts about Portuguese stress that are of interest to my topic, leaning mainly on Magalhães (2016).[6] “Whereas scholars disagree about the conditioning factors that account for the distribution of main stress in non-verbs, there is agreement that stress in verbs interacts with morphology” (Magalhães 2016: 109). That is, main stress assignment depends on the verb’s TAM features, with a triple contrast of past vs. future vs. present (Mateus & d’Andrade 2000; Wetzels 2006). “Past tense forms are accented on the last vowel of the verbal theme”; “Future tense forms are stressed on the syllable containing the TAM morpheme”; present tense forms “are stressed on the last vowel of the root”, except at 1PL and 2PL which are stressed on the theme vowel (Magalhães 2016: 109-111).

The controversy with nominals (nouns and adjectives) is about the role of syllable weight in stress assignment. According to one school of thought (e.g. Wetzel 2006) “Stress in non-verbs is predominantly prefinal in words ending in

[6] The relevant language for my purpose is of course sixteenth century (Middle) Portuguese rather than Modern Portuguese. Stress assignment principles do not seem to have changed between the two stages, however.

a vowel. When the word-final syllable ends in a diphthong or a syllable closed by a consonant, final stress is the normal case" (Magalhães 2016: 111). Whether the word ends with a light or heavy syllable is thus the conditioning factor: cf. *pato* /ˈpa•tu/ 'duck', *casaco* /kɐ•ˈsa•ku/ 'jacket' vs. *fregués* /frə•ˈgɛs/, *pomar* /pu•ˈmar/ 'orchard', *refém* /ʁə•ˈfɐ̃j/ 'hostage', *Cacheu* /kɐ•ˈʃɛw/ 'Cacheu (a town in Guinea-Bissau)'.

According to the competing account (e.g. Mateus & d'Andrade 2000), the dominant stress location in nominals is on the last syllable of the root. In *pato* and *casaco* the final vowel is a thematic vowel related to masculine gender, so that the root ends at the consonant preceding it and is stressed accordingly: [pát]$_{root}$ o]$_{theme}$], [ka•sák]$_{root}$ o]$_{theme}$]. *Fregués*, *pomar*, and *refém* are athematic nouns, therefore stressed on the final root syllable coinciding with end of word.

Nouns ending in a stressed vowel such as *avô*, *café*, *javali*, *peru*, etc. naturally fall in with this account, since they are athematic, whereas they have to be counted as exceptional in the syllable-weight account. On the other hand, proparoxytonic vowel-final nominals such as *cântara* or *óptimo* 'excellent' and paroxytonic consonant-final nominals such as *fácil* 'easy' and *açúcar* 'sugar' are exceptional for both accounts. Allowing only for two exceptional cases instead of three might be considered an asset of the root-end account.

A few remarks are in order as they are of importance for the subject matter of the present study. First, something has to be said of the nonfinite verb forms, i.e. the infinitive, the participle and the gerund. The stress pattern of the non-inflected as well as inflected infinitive seems to be the same as that of the past tense forms: *falar* [fal]$_{root}$ á]$_{theme}$ r]$_{inf}$ 'to speak', *falarmos* [fal]$_{root}$ á]$_{theme}$ r]$_{inf}$ mos]$_{1pl}$ 'for us to speak'. The participle (*falado* 'spoken') and the gerund (*falando* 'speaking'), in contrast, are clearly accented as nominals are, on the penultimate either because they end in a light syllable or because the final vowel is a theme vowel.

The second remark has to do with plural inflection in nominals. Its exponent being final *-s* in vowel-final items (*patos* 'ducks') and final *-es* in consonant-final items (*fregueses* 'customers'), plural nominals always are consonant-final and ought to bear final stress according to the syllable-weight account. A provision then has to be made to the effect that inflectional final *-s* be extrametrical, outside the stress assignment domain, so that singular and plural forms are stressed alike. The root-end account does not need this provision: cf. [pát]$_{root}$ o]$_{theme}$] s$_{pl}$], [fregués]$_{root}$ es$_{pl}$], which might be another reason for preferring it to its competitor.

But must we make a choice? Probably not, because — that is my third and last remark — we are dealing not with Portuguese as a native language, but

with pidgin / creole Portuguese, first of all with the sixteenth century Basic Variety known as Língua de Preto. That is, we are dealing with a particular situation where an L2 language had to be acquired by adults without any formal instruction — no Portuguese-as-a-second-language courses — so that they had to rely on their perception of the outer sound shape of words and their ability to reproduce it faithfully enough that they could make themselves understood.[7] In such conditions, it matters little which account of stress assignment is correct, that is a plausible simulation of the mental processes that make native speakers recognize whether an existing word is rightly or wrongly stressed and attribute stress to new words (see discussion in Magalhães 2016).[8] In fact, the real question is whether the non-native learners' native competence allowed them to hear distinctive stress in the words that struck their ears.[9] I will return to this point.

Assuming for the moment that they did hear it, what matters is the raw fact that most vowel-final words bear penultimate stress and most consonant-final words are oxytonic. Relevant exceptions — that is exceptions frequently occurring in actual speech — are mainly due to nominal plural inflection and verbal TAM inflection. As we shall see, these inflections are not present in the Basic Variety.

3. Stress in Língua de Preto

For the history of Língua de Preto and why one is justified in considering it representative of the Basic Varieties at the root of the ensuing PRCs I again refer the readers to Kihm & Rougé (2013). Let me just restate a few facts. A sizable community of West African slaves was present in Portugal from the end of the fifteenth century onwards (see Tinhorão 1988). They soon acquired an approximation to Portuguese, which they used to communicate with the European environment and among themselves, as their various L1s offered little

[7] The 'unguided' character of the learning process is a crucial feature of the situation. A recent state-of-the-art chapter about the acquisition of Portuguese as an L2 (Madeira 2016) makes it clear that all the studies commented on were about learners "in a formal instructional setting". The levels of knowledge attainable in such settings go well beyond what the Língua de Preto speakers managed to grab from the Portuguese speech surrounding them.

[8] Of course assuming there is more to it than memorizing plus analogical extension, a possibility not to be dismissed out of hand.

[9] It is well known, for instance, that owing to the absence of lexical stress in French most (European) Francophones are stress-deaf and do not hear the difference between *cântara* and *cantará*.

mutual understanding. The Basic Variety thus achieved was sufficiently institutionalized to be given a name.[10] We know it (or of it) through the Portuguese theatre of the time. African characters speaking 'Black speech' — but impersonated by masquerading white actors — appear in several plays (*autos* 'acts') mainly by Gil Vicente (1465?-1537?) and António Ribeiro Chiado (1520?-1591) (see Teyssier 1959; Berardinelli & Menegaz 1994).[11] As for the accuracy of such literary renditions, Chasca (1946: 323), in what may be the first in-depth study of an Iberian Basic Variety (Habla de Negro), opines that "in spite of considerable distortion for comic effect, Rueda and his fellows reproduced the essential peculiarities of contemporary negro speech", a conclusion also reached by Santos Morillo (2020: 14-16) as well as by Teyssier (1959) and Kihm & Rougé (2013) for Língua de Preto.

As far as suprasegmentals are concerned, we get some clue from the fact that stressed syllables are fairly regularly flagged — e.g. with a graphic accent — whenever they do not match what Portuguese readers would have expected given their native competence that makes stress notation generally unnecessary. Actually, the device appears to have been called for and used only for one kind of items, namely verbs. Here the crucial point is that verbal inflections were largely discarded in Língua de Preto — which does not mean that inflected verb forms are not encountered occasionally, but that in most cases verbs appear under one all-purpose form that does not express any particular TAM value and shows no agreement. Here is an example from Gil Vicente's *Frágoa d'Amor* ('Forge of Love') (1524, 470-474, Teyssier 1959: 234):[12]

Já mão minha branco estae,	Minha mão já está branca,	Now my hand is white,
e aqui perna branco he,	e aqui a minha perna está branca.	and my leg is white too.
Mas a mi falá Guinee,	Mas eu falo Guiné.	But I speak Guinea.
Se a mi negro falae,	Se eu falo negro,	If I speak black,
a mi branco para quê?	de que me serve ser branco?	Why be white?

A verb form such as *falá* in the third line ought to have been assigned penultimate stress by any account. Yet it is supposed to be uttered with stress on the final syllable, in a quite surprising way from a native speaker's point of view, hence the need for some graphic marking, here an acute accent. *Estae* in

[10] Língua de Preto is also known as *Falar Guiné* 'Guinea Speech', where *Guiné* refers to Sub-Saharan West Africa.

[11] In Spain Lope de Rueda (1510-1565) and Diego Sánchez de Badajoz (1479-1549) are the main playwrights to give us samples of Habla de Negro (see Santos Morillo 2020).

[12] The Portuguese translation is Teyssier's, the English one is mine.

the first line probably notates the pronunciation /eˈstaj/ with regular stress on the final diphthong, to be discussed later on. One could be tempted to see it as a deviant realization of *está* /eˈsta/ 'is', were it not for *falae* in the fourth line, which, whatever the actual pronunciation may have sounded like, clearly represents a competing device for notating unexpected final stress in a form whose actual 1SG value, like that of *falá*, is only recovered thanks to the preceding pronoun (*a mi* 'I'), while the present tense value is provided by the context.

As pointed out in the introductory section, such invariable, stress-final forms actually correspond to Portuguese infinitives having lost their final /r/. This interpretation is borne out by the occurrence in the very same texts of referentially identical verb forms that either keep the infinitive final /r/ or safeguard it by means of a paragogic vowel, so that no diacritic is necessary to signal stress location.[13] While such forms are rare in Gil Vicente's and Chiado's *autos*, they multiply in the seventeeth and eighteenth centuries, perhaps because Língua de Preto as represented on the stage had by then turned into a stereotype rather than an imitation of reality (see Kihm & Rougé 2013). That the infinitive was the verb form mainly retained in the Basic Variety for use with every TAM and person-number value — but for some scattered finitely inflected forms — is further confirmed by the Habla de Negro texts, where final *-r* is often written, whether it was actually pronounced or not in actual speech.

To repeat, unexpected stress notation is far from being systematic in sixteenth century *autos*. The same word will appear at times graphically marked for final stress (e.g. *falá*, *falaa*, *falae*), and sometimes not (*fala*), possibly in the same text. What clue do we have, then, as to the stress location in unmarked verb forms?

In order to take a closer look at the phenomenon, I examined three *autos* by Gil Vicente where Língua de Preto passages are especially important: *A Frágoa d'Amor* (see above), *A Nao d'Amores* (1527), and *O Clérigo da Beira* (1529 or 1530) (Teyssier 1959: 231-242). Altogether they include 242 Língua de Preto verb tokens representing 120 verb forms (see Appendix). Transcription variants such as *bay* ~ *bae* ~ *vae* ~ *vay*, all probably pronounced /baj/ by actual speakers and meaning 'to go', count as one verb form. Among the 120 verb forms, 43 show Portuguese finite inflections. In only 36 of the latter, however, does the inflected form convey the meaning it does in Portuguese: e.g. in *Nunca rirá ome branco* 'Never will a white man say' (P *Nunca dirá homem branco*)

[13] As a Basic Variety Língua de Preto could not be uniform by definition. Nor can we expect its literary rendition to be entirely coherent.

(Teyssier 1959: 238), *rirá*, with typical initial rhotacism (see Kihm & Rougé 2013), obviously is a 3SG indicative future form, the same as P *dirá*. In contrast, *bem ~ vem* does not mean 's/he/it comes' like its Portuguese homograph *vem*, but it is used at all persons and tenses: e.g. *Poro meu vontare ami vem / abré oyo Purutugá* {by my will I come open eye Portugal} 'From my own free will I came / to visit Portugal' (P *Por minha vontade vim / visitar Portugal*) (Teyssier 1959: 236). Such forms as *rirá* or 1/3SG subjunctive past *furtase* (P *furtasse*) 'I should rob' ought therefore to be discounted. Two non-exclusive motives explain their presence in the texts: a laudable care for realism, not surprising in Gil Vicente, since Língua de Preto was an evolving Basic Variety, so that actual speech was necessarily mixed; a not so laudable (in our view) care for comic effect.

Bae ~ vae ~ vay 'to go' (P *vai* 's/he/it goes'), *bem ~ vem* 'to come' (P *vem* 's/he/it comes'), *he* 'to be' (P *é*),[14] *tem* 'to have' (P *tem* 's/he/it has'), and *saa ~ sae ~ sá* 'to be' (P *são*?) must be retained in contrast. Interestingly, these forms are found in various PRCs: cf. Guinea-Bissau-Casamance Kriyol *bai*, *bin ~ ben*, *ten* (Rougé 1988); Lung'ie *we* (P *vai*), *vika* (P. *vem cá* 'come here'), *tê* (Maurer 2009); Papiá Kristang *bai*, *beng*, *teng* (Baxter 1988). Of special interest is the copula *saa ~ sae ~ sá* 'to be', probably from *são* 'they are' — usually written *sam* and still possibly pronounced /sã/ in the sixteenth century (Teyssier 1982: 39-40) — occurring 22 times in my corpus and to this day present in the four Gulf of Guinea creoles.

This leaves us with 84 (120 – 36) verb forms we may confidently attribute to 'pure' Língua de Preto, of which seven are clearly not related to Portuguese infinitives (e.g. *vem*), 19 are ambiguous because they end in /a/ or /e/ but are not graphically marked as having final stress, e.g. *chama* possibly representing either infinitive *chamar* /ʧaˈmar/ or 3SG indicative present *chama* /ˈʧama/ or 2SG imperative *chama!* /ˈʧama/,[15] 14 are related to Portuguese infinitives and end in /r/ followed or not by a paragogic vowel (e.g. *andar* 'to walk', *chovere* 'to rain', *buscaro* 'to look for'), and 44 are related to Portuguese infinitives, do not end in /r/, but are shown to be stressed on the final syllable,

[14] <he> for *é* 'is' was a common spelling at the time. In the *Nao d'Amores* one finds a clear example of *he* with a non-3SG value : *Se boso firalga he aqui* {if you gentlewoman be here} 'If you are a gentlewoman here' (P *Se vós sois fidalga aqui*) (l. 556) ; and an ambiguous one : *a masa firalgo he a mi* {theF more gentleman be I} 'I am the greatest gentleman' (P *o mais fidalgo sou eu*) (l. 560), as it is not clear, despite the Portuguese translation, which of the two NPs is the subject (see Teyssier 1959 : 236 for the translations).

[15] By the time of Gil Vicente's and Chiado's plays graphic <ch> still represented the affricate [ʧ].

27 by way of an acute or circumflex accent or vowel doubling or because of ending in (always stressed) /i/ (*comê* ‘to eat’, *falá* ‘to speak’, *furutaa* ‘to steal’, *dormi* ‘to sleep’), 17 owing to what looks like a final /aj/ diphthong variously spelled (*crivaninhae* ‘to scribble’, *sapantay* ‘to frighten’).

What shall we do with the 19 ambiguous cases? An interesting observation is that only seven of them end in /e/ while the remaining 12 all end in /a/ (see Appendix). Altogether my corpus of 84 ‘pure’ Língua de Preto verb forms contains no more than 17 *e*-final forms, ten of which are marked for final stress, e.g. *bebee* or *bevê* ‘to drink’ (P *beber*). Now final /e/ is a case where penultimate instead of final stress did entail a noticeable phonological difference for sixteenth-century Portuguese readers as much as it does for their present-day successors. Although not yet raised to [i] in Middle Portuguese — and still less quasi mute as in Modern European Portuguese — final unstressed /e/ was higher than stressed /e/ — “*muito fechado*” in the words of Teyssier (1982: 23, 48) — hence the need for some marking. In contrast, final unstressed /a/ was not reduced to [ɐ] as it is in Modern European Portuguese, but probably realised roughly as [a] as in Brazilian Portuguese. As a consequence, we ought not to be surprised if the authors or copyists had now and again forgotten or neglected to record final stressed /a/ because pronouncing it unstressed would not have made much difference at the segmental level. I could therefore feel authorized to add the 12 unmarked *a*-final forms to the 44 marked forms. I refrain from doing so, however, since final stress is not certain.

The provenance of the final /aj/ diphthong of *crivaninhae* ‘to scribble’ (P *escrevaninhar*), *sapantay* ‘to frighten’ (P *espantar*), etc. is unclear. Three possibilities come to mind. It could represent the now archaic 2PL imperative, e.g. *Escrevaninhai!* [eskrevaniˈɲaj] ‘Scribble!’. That part of the Língua de Preto verb stock originated from the Portuguese imperative rather than the infinitive is certainly conceivable — it is indeed one of the possible interpretations of the unmarked forms such as *chama* as we saw above — but why the plural? Another option is that the diphthong was analogically transferred from what must have been one of the very first verbs to be picked up by the incoming African slaves, namely *bae ~ bay ~ vae ~ vay* ‘to go’ (P *vai* ‘s/he/it goes’). Or, possibly the likeliest hypothesis, it was an approximate transcription of how the Portuguese perceived the Africans’ pronunciation of final stressed nuclei, perhaps variably involving some offglide — recall their linguistic background was quite diverse. The fact that only <ae> ~ <ay> is found but never <ey> — unless a transcription such as *queree* ought to be so read — may support the second option, but it is also compatible with the third one. I will return to the issue.

Be it as it may, one may retain that using the graphic form of a diphthong was one way to alert the readers to final stress. The 17 verb forms ending in <ae> or <ay> may therefore be merged with the 27 marked forms, hence the 44 overtly stress-final forms mentioned above representing 52 per cent of the 84 'pure' Língua de Preto verb forms. The 40 remaining forms either are not related at all to Portuguese infinitives (7), or are graphically unmarked (19), or do not end in a stressed light syllable but in <r> (8), <re> (4), or <ro> (2) (see Appendix). Adding the eight *r*-final forms, since they obviously bear final stress as well, increases the amount to 52, i.e. 62 per cent.

The percentage of overtly stress-final items appears significantly less in terms of tokens, namely 37.20 per cent of the 242 verb occurrences. This is due to the high frequency of a few (apparently) inflected forms: *bae* and variants 'to go' (16 tokens), *disse* and variants 'he said' (9 tokens), *he* 'to be' (15 tokens), *sa* and variants 'to be' (22 tokens), accounting for 25.62 per cent of all occurrences.

Now, as already pointed to, the 44 verb forms showing a stressed final light syllable (including graphic diphthongs that may not corrrespond to real diphthongs) are of especial interest in that they do not abide by any conceivable Portuguese stress rule. And the cause of their peculiarity is precisely their having discarded the final /r/ of the Portuguese infinitives they are related to, which raises the issue of word-final consonants in Portuguese and their fate in Língua de Preto.

The inventory of word-final consonants is limited in Portuguese: only the sonorants /l/, /n/ and /r/ and the obstruents /s/ and /z/ are legitimate.[16] (Recall that final /e/ was never mute in Middle Portuguese, so that *leque* 'fan' was pronounced ['lɛke], not [lɛk] as it may be in some varieties of Modern European Portuguese.) In Língua de Preto final /z/ — infrequent and still realised [z] or [s] in the sixteenth and early seventeenth century (see Teyssier 1982: 46-47) — was never deleted: e.g. *nariz* 'nose' (Teyssier 1959: 234). Taking into account that plural *-s* is foreign to Língua de Preto, where nominals show no number inflection at all, one finds a few instances of final /s/ deletion when it either is not a morph as in *Tordesilla* for *Tordesillas*, or is a morph that probably was not recognized as such by the Língua de Preto speakers as in *vamo* 'let's go' for *vamos*. Teyssier (1959: 243) points out one case – but a momentous one! – of

[16] Final /n/ — written <m> or as an accent (*til*) or <n> in a few learned or foreign words such as *hífen* 'hyphen' — never appears as such, but only by nasalizing the preceding vowel or diphthong.

final /l/ deletion in *Purutugá* for *Portugal*. (In Kriyol *tuga* /ˈtuga/ still means ‘European’ or ‘Portuguese’.)

Final consonant deletion therefore mainly concerns /r/, yet with a catch as already mentioned: it is never deleted in nouns and adjectives, being either kept as such: e.g. *amor* ‘love’, *mar* ‘sea’, *mior* ‘better’ (P *melhor*), *moyer* ‘woman’ (P *mulher*), etc.; or protected by a paragogic vowel according to a general trend towards open syllabification: e.g. *amoro* ‘love’ (P *amor*), *muyere* ‘woman’ (P *mulher*), *seoro* ‘sir’ (P *senhor*), etc.

We thus come to the basic question of the present study: Why did final /r/ only dropped in verb forms proceeding from Portuguese infinitives? Before trying to answer this question, let us have a look at the PRCs in order to ascertain whether what happened in Língua de Preto represents a crucial change with enduring consequences.

4. Suprasegmentals of prominence in the PRCs

PRCs divide into two groups as far as suprasegmentals are concerned according to whether they developed tone systems or remained stress languages like the lexifier. The four Gulf of Guinea PRCs and Papiamentu belong to the first group, Upper Guinea PRCs (Guinea-Bissau-Casamance Kriyol and Cape Verdean) and Asian PRCs to the second. Contact readily explains the divide. Only Atlantic PRCs had a chance to transfer tone from the Bantu, Edoid and Kwa Niger-Congo languages they were in contact with (see e.g. Hagemeijer & Ogie 2011). In contrast, Indo-Portuguese, Sri-Lankan Portuguese, Papiá Kristang, and Batavia Creole always were in contact with Indo-Aryan, Dravidian and Austronesian languages where tone plays no grammatical role. Being themselves related to a non tonal language, one does not see why and how they would have developed a tone system. Only Macao Creole Portuguese (aka Macanese or Patuá) was — and is, to the extent it is still spoken — in contact with a tone language, Cantonese. Yet Macanese is widely recognized to be an offshoot of Malayo-Portuguese, whose phonology was already stabilized when it came to Macao and unlikely to be influenced by the highly complex tone system of Cantonese (see Holm 1989: 296-298). The case of Upper Guinea PRCs is more complex, as they were (and on the continent still are) in contact with tone as well as stress languages. The latter apparently were more influential in the prosodic domain (see below).

An all-encompassing change was the disappearance of gender (or word class) as a grammatical category, probably as soon as the Basic Variety stage

(see Kihm & Rougé 2013). It entailed that the final vowels in nominals such as *casa* 'house', *amigo* 'friend', *ventre* 'belly', *grande* 'big', *formoso* 'beautiful', etc. were absorbed into the root when they were not dropped altogether (Luís 2008), so that only information from the surface syllabic structure remained accessible in any event.

4.1. Guinea-Bissau-Casamance Kriyol[17]

Guinea-Bissau-Casamance Kriyol is very much like Língua de Preto except for systematicity. In nominals (nouns and adjectives) stress is located according to syllable structure, penultimate if the final syllable is light, final otherwise: e.g. *kasa* /ˈka•sa/ 'house', *amigu* /a•ˈmi•gu/ 'friend', *kanibeti* /ka•ni•ˈbe•ti/ 'penknife' (P *canivete*), *katchur* /ka•ˈʧur/ (B) ~ /ka•ˈcur/ (Z) 'dog' (P *cachorro* 'puppy'), *pilon* /pi•ˈlõŋ/ 'mortar' (P *pilão* 'pestle'), *burmedju* /bur•ˈme•ʤu/ (B) ~ /bur•ˈme•ɟu/ (Z) 'red' (P *vermelho*), *ferus* /fe•ˈrus/ 'ferrous' (P *ferroso*).[18] In verbs, stress is located on the final light syllable: e.g. *kasá* 'to marry' /ka•ˈsa/ (P *casar(-se)*), *kumé* /ku•ˈme/, 'to eat' (P *comer*), *diskisí* /dis•ki•ˈsi/ 'to forget' (P *esquecer*). This extends to verbs borrowed from endemic languages, even when not ending in /a/, /e/, or /i/, e.g. *bambu* /bã•ˈmbu/ 'to carry on one's back', *djubi* /ʤuˈbi/ (B) ~ *djobe* /ɟoˈbe/ (Z) 'to look', *potcholi* /poʧoˈli/(B) ~ /pocoˈli/ (Z) 'to crush' (Kihm 1994; Chapouto 2014: 17-18; Biagui 2017). Note that, although the placement rule may have changed — depending on the account one retains for Portuguese stress placement — the position of stress remained entirely stable.[19]

There are exceptions with nominals, actually the same as in the lexifier. Some are predictable, for instance items involving the ending *-iku*, always stressed on the pre-penult: e.g. *demokrátiku* 'democratic' (P *democrático*);

[17] Biagui (2017) and Quint & Biagui (2013) consider what they call 'Casamancese Creole' a separate language from Guinea-Bissau Kriyol. They may be right as far as the native speakers' spontaneous representations are concerned. Yet both languages are entirely mutually intelligible and grammatically near identical. They shared a common history until at least 1886 when Casamance passed under French rule — but the border never was a barrier to exchanges, as witnessed by the numerous Guineans living in the regional capital Ziguinchor. For all these reasons, I follow Rougé (1994, 2004) in viewing Guinea-Bissau-Casamance Kriyol as one language with several dialectal varieties, Bissau (B) and Ziguinchor (Z) the two main ones.

[18] In words of Portuguese origin the only legitimate codas are the sonorants /r/, /l/, and /n/ (realized homorganic to a following consonant or [ŋ] word-finally), the obstruent /s/, and the approximants /j/ and /w/. All consonants are possible in words borrowed from endemic languages.

[19] This is strongly reminiscent of what is observed in the transition from Late Latin to the Romance languages.

others not: e.g. *papé* ‘father’ < *papá(i)* (see Rougé 2004 : 223 for the etymology), *mamé* ‘mother’ (P *mamã*), *lápis* ‘pencil’ (P *lápis*), *súkar* ‘sugar’ (P *açúcar*). (The latter is regularized as *sukru* in Casamance, perhaps under the influence of French *sucre* /sykʁˢ/.)

Biagui (2017 : 27-28) takes exception to this analysis first proposed by Doneux (1979). He points out that the Casamance variety includes many words with penultimate stress even though they end with a consonant: cf. *pálum* ‘palm (of hand)’, *pádur* ‘priest’, *útur* ‘other’, etc. Such forms are easily accounted for, however, if one looks at their Guinea-Bissau counterparts *palmu*, *padri*, *utru*: in Casamance the inherited final vowels (cf. *palmo*, *padre*, *outro*) dropped, producing a final consonant cluster that was broken by epenthetic /u/. One sees here a rather regular correspondence between the two varieties: in disyllabic nouns whose Portuguese etyma show the pattern /CV•CC$_{sonorant}$V/ (*padre*) or /CVC•C$_{sonorant}$V/ (*palmo*), the pattern was preserved in Guinea-Bissau, but it changed to /CV•CuC$_{sonorant}$/ in Casamance.

As a consequence, stress, from regular that it is in Guinea-Bissau (penultimate in a vowel-final noun), became irregular in Casamance in these words. Exceptions do not prove the rule, as the saying goes, but they are evidence that there is a rule. Moreover, such forms may actually be brought back to the fold by capitalizing on the fact that they all involve a sonorant. It can therefore be assumed that final /m/ in e.g. *pálum* is syllabic, /u/ realising the vocalic release required for actual pronunciation. With /CV•CV$_{sonorant[+syl}$/ as a syllable structure, stress placement on the first syllable would abide by the general rule in these items. Note that nouns borrowed from the endemic languages follow the general rule: cf. *badjúda* ‘young girl’, *mungút* ‘hornless black cow’, *badjík* ‘wild spinach’, the first and the last ones possibly from the Atlantic language Ñun (also called Baynuk) (Rougé 2004: 292), etc.

As in the lexifier, stress placement in nominals is not sensitive to inflection. The only nominal inflection in Guinea-Bissau-Casamance Kriyol is pluralization by way of a /-(i/u)s/ suffix depending on whether the root ends in a vowel or a consonant: cf. *amigu* /aˈmi•gu/ vs. *amigus* /aˈmi•gu-s/ ‘friend(s)’, *katchur* /kaˈʧur/ (B) ~ /kaˈcur/ (Z) vs. *katchuris* /kaˈʧu•r-is/ (B) ~ *katchurus* /kaˈcu•r-us/ (Z) ‘dog(s)’. Although the words now end in a consonant, stress stays on the same syllable as in the singular. Since there is no empirical evidence for distinguishing roots and stems in Guinea-Bissau-Casamance Kriyol, extrametricality of plural /s/ looks like the best explanation, with /i ~ u/ an epenthetic vowel called for by a general ban on final consonant clusters.

Verbs only inflect synthetically for passive voice, by way of a /-du/ suffix that has no effect on stress placement: cf. *kumé* /kuˈme/ ‘eat’ vs. *kumedu*

/ku•ˈme•-du/ ‘be eaten’.[20] Recall that Portuguese participles, the obvious origin of Kriyol passive forms, are stressed the nominal way. In Guinea-Bissau-Casamance Kriyol, however, passive forms have nothing of the mixed character of participles, they are strictly verbal, so that extrametricality again recommends itself.

Given this, regular stress assignment in Guinea-Bissau-Casamance Kriyol may be formalized as the two following metarules:

(1) $W_N \Leftrightarrow$ (i) $\langle \sigma^{*}\acute{\sigma}\varsigma \rangle \vee$ (ii) $\langle \sigma^{+}\acute{\Sigma} \rangle$

(2) $W_V \Leftrightarrow \langle \sigma^{+}\acute{\varsigma} \rangle$

Metarules (1) and (2) are generalizations over noun word-forms (W_N) and verb word-forms (W_V) realizing nominal and verbal lexemes. Word forms are represented as strings of syllables notated as sigmas. Upper case sigma (Σ) notates heavy (CVC) syllables, final lower-case sigma (ς) notates light CV syllables, non-final lower-case sigma (σ) notates both types. Stressed syllables are accented. In (1) pre-stress syllables are assigned a Kleene star in case (i) as their cardinality ranges from 0 to *n*, the usual value being 0, yielding disyllabic noun forms, more rarely 1 or 2. They are assigned a Kleene plus in case (ii) of (1) and in (2) since there must be at least one pre-stress syllable to ensure a stress contrast in the word. Rules (1) and (2) are couched as biconditionals: if a word is nominal or verbal, then it shows one of the stress patterns described on the right of the double arrows; if it shows one of these stress patterns, then it is nominal or verbal.[21]

Papé, *súkar*, etc. must be lexically marked as exceptions to (1). Monosyllabic nouns, light (*po* ‘stick, tree’, *cha* ‘tea’, etc.) or heavy (*kau* ‘place’, *tchur* ‘funeral’, etc.), are not catered to by (1) either, since they may be considered to be stressed by default, assuming that every free (non-clitic) form must bear one main stress.[22] Monosyllabic verbs constitute a small and closed set, and they are exceptional with respect to (2) in that their unique syllable is heavy, having /j/ or /n/ as a coda: *bai* ‘to go’, *bin* ~ *ben* ‘to come’, *den* ‘to ache’,

[20] Causitivization as in *kumente* /ku•me•ˈnte/ ‘to make eat, to feed’ pertains to derivation rather than to inflection (Kihm 1994, Biagui 2017: 203-213).

[21] Rules such as (1) and (2) are meant to be akin to the Feature Cooccurrence Restrictions (FCR) of Generalized Phrase Structure Grammar (Gazdar et al. 1985: 27-29).

[22] For reasons I cannot enter into here I analyse /aw/ as a branching rhyme (nucleus plus coda) rather than as a diphthong. The same analysis applies to /aj/, /uj/, /ej/, /ew/, and /iw/.

kai ‘to fall’, *lei* ‘to read’, *pui* ‘to put’, *sai* ‘to go out’, *tchai* ‘to be adulterous’, *ten* ‘there is/are’. [23]

Stress in Guinea-Bissau-Casamance Kriyol thus appears to be not only phonologically assigned, depending on syllable structure as per rule (1), but also lexically assigned, being determined by the POS a given lexeme belongs to.

Verb stress as described by (2) is not entirely fixed insofar as the rule applies to isolated verb forms. In disyllabic verbs such as *kanta* ‘to sing’ a stress is heard on the initial syllable of the perfective base verb form at 1SG: *N kanta* /ˈŋkã•nta/ ‘I sang’. Andrade & Kihm (2000: 104) account for this raising by assuming that procliticizing *N* ‘I’ yields a complex, heavy onset that attracts stress. On the other hand, in words of more than two syllables, no matter whether verbs or nominals, the initial syllable bears secundary stress: e.g. *karmusa* /ˌkarˈmusa/ ‘ceremonial dance’, *kurinti* /ˌkurinˈti/ ‘to drive’ (Bull 1989: 77; Chapouto 2014: 17-18). Insofar as *N* is part of the phonological word and usually syllabic, it may well be that that initial stress in 1SG base verb forms actually is secundary stress due to phonological word extension, so that a more accurate transcription of *N kanta* should be /ŋ•kã•ˈnta/. Guinea-Bissau-Casamance Kriyol phonology has not been sufficiently explored yet to decide on this point.[24]

It is only fair to mention that the first author to propose a scientific description of Guinea-Bissau-Casamance Kriyol, Wilson (1962: 11), reaches a different conclusion. According to him, disyllabic verbs and some trisyllabic verbs get stress on the first syllable, which stress moves to the final syllable only when enclitic object pronouns attach to the verb as in *i kanta-l* /ikãnˈtal/ ‘s/he/it sang it’. All later studies agree that main stress in verbs is final in contrast with main stress in nominals, hence a number of minimal pairs such as *koba* /ˈko•ba/ ‘hole’ vs. *koba* /ko•ˈba/ ‘to dig’ (Biagui 2018: 23). Perhaps Wilson, excellent field linguist that he was, was led astray by the secundary stress phenomenon just pointed to.

4.2. Cape Verdean

Both dialect groups making up Cape Verdean have stress as their only suprasegmental of prominence, just like Guinea-Bissau-Casamance Kriyol.

[23] Except *tchai*, of uncertain origin (Rougé 2004:356-357), these verbs all come from the 3SG present indicative of irregular Portuguese verbs and at least three of them can be traced back to Língua de Preto (see above).

[24] There is evidence that the other clitic subject pronouns (*bu*, *i*, *no*, *bo*, *e*) do not form one phonological word with the succeeding verb (Rougé & Kihm 2018).

They differ, however, insofar as stress placement is concerned. In the Barlavento group (São Vicente, Santo Antão, São Nicolau) as in Kriyol stress placement depends primarily on POS: verbs are stressed on the final vowel: cf. *txora* /ʧo•ˈra/ 'to cry', *forma* /for•ˈma/ 'to form', *bombú* /bɔ•ˈmbu/ 'to carry on one's back'; nominals are stressed on the final syllable if it is heavy, on the penultimate otherwise: cf. *amig* /a•ˈmig/ 'friend', *forma* /ˈfɔr•mɐ/ 'form' (Cardoso 1989: 96; Andrade & Kihm 2000; Delgado 2008: 103ff.; Swolkien 2013: 23). Barlavento Cape Verdean thus abides by (1) and (2). The main difference between it and Guinea-Bissau-Casamance Kriyol is that, owing to deletion of final unstressed vowels in nouns (cf. *amig* above) the inventory of final heavy syllables is much larger than it is in Guinea-Bissau-Casamance Kriyol.

Sotavento Cape Verdean is not so homogeneous. Of its three main varieties, Santiago, Fogo, and Brava, only Fogo is like Barlavento and Guinea-Bissau-Casamance Kriyol in assigning stress according to POS (Tavares Moreira 2020: 108-114). In contrast, Santiago and Brava have syllable weight as the determinant of stress placement across the board: exceptions apart, words ending in a light syllable have penultimate stress; words ending in a heavy syllable, with the same constraints on possible codas as in Guinea-Bissau-Casamance Kriyol, have final stress (Meintel 1975; Veiga 1982: 62; Andrade & Kihm 2000; Lang 2013: 5). That this pattern may not reflect the initial state of affairs, however, but results from later regularization is suggested by an intriguing phenomenon pointed out by Lang (2013: 5).

In noun-verb pairs where stress falls on a non-high vowel, the vowel is lowered in the noun, raised in the verb: cf. *karéka* /ka•ˈrɛ•kɐ/ 'bald head' vs. *kareka* /ka•ˈre•kɐ/ 'to become bald', *karapáti* /ka•ra•ˈpa•ti/ 'tick' vs. *karapati* /ka•ra•ˈpɐ•ti/ 'to cling to', *frónta* /ˈfrɔ̃•ntɐ/ 'misfortune' vs. *fronta* /ˈfrõ•ntɐ/ 'to suffer a misfortune'. A historical explanation related to differences in stress placement seems indeed feasible. Assume that at an earlier stage Santiago Cape Verdean was like Guinea-Bissau-Casamance Kriyol, Fogo and the Barlavento variety in having preserved the Portuguese and Língua de Preto stress pattern, so that verbs were stressed on the last syllable like the etymological Portuguese infinitives. In that configuration raising of pre-stress non-high vowels was the rule in Middle Portuguese, so that verbal **fronta* < *afrontar* would have been pronounced /frõ•ˈnta/ — compare Guinea-Bissau-Casamance Kriyol *f(o)ronta* /f(o)(•)rõ•ˈnta/, although the height contrast is no longer phonemic in Kriyol — whereas nominal **fronta* < *afronte* stressed on /ɔ̃/ was already /ˈfrɔ̃•ntɐ/. When stress placement was delexicalized at a later stage, the height contrast remained as a categorial clue.

4.3. Papiá Kristang

Papiá Kristang as described in Baxter (1988, 2013) is proof that stress placement in PRCs is insensitive to the East-West divide. As in Guinea-Bissau-Casamance Kriyol and Barlavanto and Fogo Cape Verdean stress location depends on POS: verbs end in vowels and have final stress (e.g. *kumí* 'to eat'), whereas, but for a handful of exceptions, vowel-final nouns have penultimate stress (e.g. *káza* 'house') (Baxter 1988: 37-38). Moreover, as in Guinea-Bissau-Casamance Kriyol and Barlavento and Fogo Cape Verdean, the immediate cause of the verbs' final stress, /r/ deletion, only affected Portuguese infinitives: cf. *mar* 'sea', *dibinyador* 'soothsayer', etc.

4.4. Gulf of Guinea Creoles

Gulf of Guinea Creoles constitute a group of closely related languages including Santome and Angolar on the island of São Tomé, Lung'ie on the island of Príncipe, and Fa d'Ambô on the island of Annobón (Equatorial Guinea). Lung'ie phonology enjoys the best description thanks to Traills & Ferraz (1981), Maurer (2009: 14-27), Agostinho (2015), and Agostinho & Hyman (2021). I will rely on the second and the last.

4.4.1. Lung'ie

According to Maurer (2009), Lung'ie is a tone language with a fully specified system based on the contrast of a high (H) and a low (L) tone. The existence of a few monosyllables only contrasting for tone supports this claim. Perusing Maurer's (2009: 211ff.) word-list one finds such minimal pairs as, for instance, *txi* /ʧi^{H}/ 'uncle' (P *tio*) vs. *txi* /ʧi^{L}/ 'you$_{sg}$' (P *ti*), *txya* /ʧja^{H}/ 'aunt' (P *tia*) vs. *txya* /ʧja^{L}/ 'to take out' (P *tirar*), *va* /va^{H}/ 'whip, board used in the construction of a roof' vs. *va* /va^{L}/ 'split, cut, chop, tear'. Such pairs seem to be rare, but it is enough that they exist.

Agostinho & Hyman (2021) agree with Maurer's facts, but they propose a different interpretation. According to them Lung'ie shows a privative contrast H vs. no tone (∅) and there can be only one H in a phonological word. A word like *káxi* 'house' to which Maurer assigns a HH pattern (/ka^{H}•ʃi^{H}/), they therefore analyse as /ka^{H}•ʃi$^{∅}$/, where only the first syllable bears phonemic tone, and they account for the high pitch of the second syllable by a "spilling over" of the pitch of the first syllable. In *kasó* 'dog', LH for Maurer, the low pitch of the first syllable represents the default realization of ∅ for Agostinho & Hyman,

who regard monosyllables like *txya* 'to take out' and plurisyllables like *ôtô* 'neck' as toneless rather than low.[25]

Given the purpose of the present study, I feel under no obligation to side with one or the other account. Assuming a privative opposition rather than full specifying (equipollent opposition in Troubetzkoy's terms) is certainly justified whenever a given feature shows only two values, one of which may be viewed as the mere absence of the other as in voiceless vs. voiced. What is important for my purpose is that both authors agree on the following two points: (i) the active suprasegmental in Lung'ie is tone rather than stress;[26] (ii) for words with Portuguese etyma tone assignment depends on POS.

Indeed, according to Maurer (2009: 26), "[g]rammatical tone is used to distinguish lexical classes […] in most cases the noun has a HH pattern and the verb a LL one": cf. *fala* /fa^{H}•la^{H}/ 'speech, talk' vs. *fala* /fa^{L}•la^{L}/ 'to speak'. For Agostinho & Hyman (2021: 20) "[v]erbs are toneless with a few exceptions" (whereas nouns bear one H), so that *fala* 'to speak' must be represented as /fa$^{\emptyset}$•la$^{\emptyset}$/. This pattern is reproduced in monosyllables as shown by the above examples of *txya* 'aunt' (H) vs. *txya* 'to take out' (L/∅) and *va* 'whip' (H) vs. *va* 'split' (L/∅). And note that verbal *va*'s etymon is given as Edo *vá* 'break into parts' whose tone is high, which suggests that the Lung'ie lexical pattern won over etymology. (No etymon is given for *va* 'whip'.) In other words, what the languages in the foregoing sections do with stress, Lung'ie does with tone.

There are two puzzles, however. One is raised by the tonelessness (or uniform low tone) of verbs. If verbs come from Portuguese *r*-less infinitives as they probably do, and given the stress-high tone diachronic correspondence, one expects to find a H on the final syllable, contrary to fact. Agostinho & Hyman offer no explanation, nor do Traill & Ferraz (1981). A clue may come from Maurer's (2009: 14) remark that "words in isolation are treated as whole utterances, in the sense that they start on a mid or high pitch and end on a low pitch". As a matter of fact, except yes-no questions not involving the clause-final low tone (or toneless) question particle *a* (Maurer 2013b: 78), Lung'ie

[25] Lung'ie tones are register tones whose pitch does not vary during production (Hyman 1975: 214). Contour tones are heard, but they only occur with long vowels which Maurer and Agostinho & Hyman analyse as sequences of identical vowels, each bearing its own register tone (Maurer) or one bearing the lone H allowed in a word (Agostinho & Hyman): cf. *kôôsu* /ko^{L}o^{H}su^{H}/ or /ko$^{\emptyset}$o^{H}su$^{\emptyset}$/ 'stone (of fruit)' realized [ko:^{LH}su^{H}].

[26] Maurer assumes stress to be active as well: "The stressed syllables in Principense seem to correspond to the original Portuguese stress" (p. 26). Agostinho & Hyman point out that since Portuguese stress has been almost invariably reinterpreted as high tone and there can be only one H per word — probably a consequence of this reinterpretation — assuming stress becomes unnecessary.

utterances always end on a phonetic low tone owing to downdrift. For instance, in *Pôkô vika* 'The pig came', the second (or sole) H of *vika*, an exceptionally H-final verb, is actually realized low (Maurer 2009: 25-26). Now verbs, intransitive or intransitively used, are especially prone to occur utterance-finally, in such a way that an original final high tone resulting from Portuguese final stress would rarely have surfaced as such, but would have been lowered most of the time. Could this be a plausible diachronic explanation of the generalization of tonelessness or a LL pattern to all verbs but a few, resulting in an overall tonal contrast between verbs and nouns?

The second puzzle, which Agostinho & Hyman call a "mystery", is the fact that a large majority (70 per cent) of nouns of African (mostly Edoid) provenance are toneless (or all low). In Maurer's (2009) word list I found 10 words provided with an Edoid etymology, yet showing high tones: e.g. *aké* /a^{L}•kɛ^{H}/ 'clay pot used for water' (Edo *àxé*), *ikôkô* /i^{H}•ko^{L}•ko^{L}/ 'coco yam, taro' (Edo *íjòxó*, Emai *íxùòxúó*), *kukundya* /ku^{L}•kun^{H}•djá^{H}/ 'coconut' (Edo *èkòkódjà*), *kyêwu* /kje^{H}•wu^{H}/ 'stone of the fireplace' (Edo *ìkéwú* 'three stones serving as pot rest', etc. Interestingly these words are all nouns but for one exception: *bi* /bi^{H}/ 'to push (unintentionally' (Edo *bí* 'to push, to shove). Toneless / all low items, in contrast, include many verbs such as *baa* 'to burn, to shine' (Edo *bàá* 'to be red, to shine'), *dumu* 'to grind' (Edo *dụ̀mụ́*), *fu* 'to clean (with soap)' (Edo *hú* 'foam, to form a lot of suds'), etc.

Apparently, this observation can be generalized, for Agostinho & Hyman (2021: 20, Table 15) give 0 as the number of verbs of African origin showing a high tone — which is not quite accurate if Maurer is correct in assigning a H to *bi*, but must be very near the truth in any event. Now this is crucial, as it strongly suggests that tone assignment according to POS — H to nouns, L or no tone to verbs — was indeed the primary process that affected not only the Portuguese inherited lexicon as indicated above, but also the 20 per cent words with an identified or probable African origin (Maurer 2009: 211).

Taking stock of the outcome of this process I therefore propose the following metarule for nouns, assuming in accordance with Agostinho & Hyman that only H need be marked:

(3) $W_N \Leftrightarrow \langle \sigma^* \sigma^H \sigma^* \rangle$

Given the syllable structure of Lung'ie, σ means V, CV, or CVV. Another metarule will account for verbs:

(4) $W_V \Leftrightarrow \langle \sigma^+ \rangle$

Both rules allow for (rare) exceptions.

4.4.2. Santome and Angolar

Disagreeing with Ferraz (1979) and Schang (2003), Maurer (2008) gives evidence that Santome also is a tone language and shows the same contrast as does Lung'ie between high nouns and low verbs: "Most disyllabic verbs are to be interpreted as LL (or as tonologically neutral), and where there is a corresponding noun, this noun shows a HH tonal pattern" (Maurer 2008: 258-259). In Angolar as well "Tonal distinctions are also used to distinguish syntactic categories, especially verbs from nouns, as in *ðùrà* 'to help' vs. *ðúrá* 'help'" (Maurer 2013a), i.e. all low verbs vs. high nouns.

4.4.3. Fa d'Ambô

The most recent and complete description (Hagemeijer et al. 2020) is unfortunately a bit short on phonology. In contrast with previous studies which either did not broach or did not conclude on the issue (Barrena 1957; Ferraz 1984; Granda 1990; Zamora Segorbe 2010; Post 2013), the authors consider Fa d'Ambô a tone language like the other Gulf of Guinea Creoles, with again a contrast between nouns and verbs. But for sandhi phenomena, monosyllabic nouns bear high tone (e.g. *dô* /do^{H}/ 'pain'), whereas disyllabic nouns "show the four logical tone melodies that a two-tone system displays: HH, HL, LH, and LL" (Hagemeijer et al. 2020: 17) when followed by the high tone TAM markers *kha* /xa^{H}/ (habitual) or *skha* /sxa^{H}/ (progressive), and trisyllabic nouns exhibit six patterns (instead of the logical eight): HHL, HLL, LHL, LLL, LLH, and LHH, that is the same as with disyllables, only with an additional L for the first four ones, and an additional H for the last two. Verbs, in contrast, turn out to be uniformly low whatever their length as in Lung'ie, thus standing out from the tonal diversity of nouns. As in Lung'ie also, there are a few exceptions. It may be to the point, however, to notice that the one cited by the authors, *mendu* /mɛ̃^{H}du^{H}/ 'to fear, to be afraid' is also a noun *mendu* /mɛ̃^{H}du^{H}/ 'fear'.

As for the presence of stress along with tone, Granda's (1990: 31) assessment is cautious: "*considero, provisionalmente al menos, que el criollo annobónes actual es una lengua intensiva y no tonal*".[27] Nevertheless he points out a few minimal pairs differing tonally although not for the position of stress

[27] Granda justifies his "*provisionalmente*" by pointing out in a footnote that he was not able to perform instrumental analysis of his recordings owing to "*la carencia de aparatos adecuados*" in Madrid.

such as *patu* /ˈpa^{L}tu^{L}/ ‘dish’ (P *prato*) vs. *patu* /ˈpa^{H}t^{L}/ ‘bird’ (P *pássaro*), *kese* /ke^{L}ˈse^{H}/ ‘to forget’ (P *esquecer*) vs. *kese* /ke^{L}ˈse^{L}/ ‘to grow’ (P *crescer*), and *petu* [ˈpɛ^{L}tu^{L}] ‘black’ (P *preto*) vs. *petu* [ˈpɛ^{H}tu^{L}] ‘near’ (P *perto*). In Hagemeijer et al. (2020), these pairs are given as *paatu* [LLH] ‘dish’ vs. *paatu* [HHL] ‘bird’, *kêsê* [LL] ‘to forget’ vs. *kêêsê* [LLL] ‘to grow’, and *peetu* [HHL] ‘black’ or ‘near’, with no mention of stress.[28] Judging by the word lists, there seems to be some correlation in nouns between high tone in the Fa d’Ambô words and stress in the Portuguese or Spanish etyma, especially in longer words: e.g. *abwela* /a^{L}bwɛ^{H}la^{L}/ ‘grandmother’ (S *abuela*), *batxizamentu* /ba^{L}txi^{L}za^{L}mɛ̃^{H}tu^{H}/ ‘baptism’ (P *batizamento*), *ladalan* /la^{L}da^{L}lã^{H}/ ‘thief’ (P *ladrão*), *lavulu* /la^{H}vu^{L}lu^{L}/ ‘book’ (P *livro*), etc. Whether stress is realized in addition to tone is not said. It is quite possible that the stress-high tone correlation is a purely diachronic one as in Lung’ie.

4.4.4. Papiamentu

Stress and tone interact in Papiamentu. I can do no better, I think, than referring to Kouwenberg (2004) who convincingly argues that “Treating tone as a grammatical marker of word category allows for a straightforward account of Pp [Papiamentu] tone placement” (p. 69). Summarizing her results, nominals show regular coincidence of stress and high tone as well as the typical Iberian surface pattern (irrespective of theoretical accounts): penultimate stress and H if the last syllable is light: e.g. *buraku* /bu•ˈra^{H}•ku/ ‘hole’ (P/S *buraco*); final stress and H if the last syllable is heavy: e.g. *fabor* /fa•ˈbor^{H}/ ‘favour’ (P/S *favor*). In verbs, all ending in a vowel, in contrast, stress and tone placement depend on syllabicity and may not coincide. In disyllabic verbs stress is initial, H final: e.g. *kaska* /ˈkas•ka^{H}/ ‘to peel’; in longer verbs stress and H meet on the final light syllable: e.g. *kuminsá* /ku^{H}•min•ˈsa^{H}/ ‘to begin’ (S *comenzar*).[29]

As a consequence nominals ending in a light syllable bearing stress and H, of which there are a few, must be counted as irregular. Deletion of a final syllable or consonant is the usual cause of the irregularity: e.g. *piská* /pi•ˈska^{H}/ ‘fish’ (S *pescado* /pes•ˈka•ðo/), *muhé* /mu•ˈhe/ ‘woman’ (S *mujer* /mu•ˈχer/),

[28] Analysing long vowels as sequences helps understand the tone pattern of *paatu* ‘bird’, taking the Portuguese etymon to be *pássaro*, with the following putative change chain where high tone and etymological stress finally coincide: /ˈpasaru/ > */ˈpasru/ > /ˈpastru/ (cf. Kriyol *pastru* ‘bird’) > */ˈpasatru/ > /pa^{H}a^{H}tu^{L}/.

[29] H on the first syllable is due to a prosodic constraint to the effect that successive tones should not be identical. Papiamentu rather clearly supports Agostinho & Hyman’s hypothesis for Lung’ie that only H is the active tone. I therefore abstain from notating low tones (also see Remijsen & Heuven 2005).

etc. Likewise verbs not showing a H on the final light syllable are erratic and generally non-Iberian: e.g. *skeiru* /ˈskɛjᴴ•ru/ ‘to brush’ (Dutch *scheren* ‘to shear’?).

4.5. Indo-Portuguese and Sri Lanka Portuguese

The PRCs of India and Sri Lanka are purely stress languages, which comes as no surprise since creolized Portuguese never was in contact with tone languages in this area. Unlike the languages examined so far, however, prominence assignment is not sensitive to POS, but it only depends on phonological factors — but Sri Lanka Portuguese might be in need of some revisiting as we shall see.

4.5.1. Diu Indo-Portuguese

Determining stress placement in Diu Indo-Portuguese is simplicity itself: “Stress falls predictably on the word’s final syllable” (Cardoso 2009a: 100). Yet note that in nominals such final stress is due to across-the-board deletion of the thematic vowel of the Portuguese etymon as in *kubert* /kuˈbɛrt/ ‘blanket’ (P *coberta*) or even of the whole post-stress syllable as in *muz* ‘music’ (P *música*). In other words, like Kriyol, Cape Verdean, and Papiá Kristang, Diu Indo-Portuguese kept stress in place in nominals. Where it differs from these languages is in verbs, in which not only did Portuguese infinitive /r/ drop, but the stressed thematic vowel as well, hence verb forms like *fal* ‘speak(s)’. Whenever inflectional suffixes are added — which only happens with verbs — stress automatically moves to the new final syllable: cf. *falo* /fa•ˈlo/ ‘spoke’, *apaŋ* / a•ˈpaŋ/ ‘get(s)’ vs. *apəŋo* / a•pə•ˈŋo/ ‘got’. Note, however, that this /-o/ ending originates in the 3SG preterite *-ou* ending of Portuguese Inflection Class I verbs, which is stressed: e.g. *falou* ‘s/he spoke’, /fɐ•ˈlo/ in Modern Portuguese, [fa•ˈlow] in Middle Portuguese. There is thus a historical explanation for final stress in such forms, even though a declaration to the effect that stress falls on the last syllable of phonological words is synchronically adequate and sufficient.

4.5.2. Korlai

Stress assignment proceeds similarly in Korlai insofar as Korlai has a “general word-final stress rule” (Clements1996: 86; also see Wetzel & Hermans 2014), which knows one exception: owing to contact with Marathi, trisyllables ending in a vowel have stress on the penult: e.g. *kombarsa* /kom•ˈbar•sa/ ‘to talk’, hence Clements’s (1996: 88) stress rule: “Stress the final syllable of the word, unless

the word is trisyllabic with open final syllable, in which case stress the penult." POS seems to enter for nothing.

Yet there is a phonological difference between nominals and verbs. In the former, as in Diu, the thematic vowels /a/, /o/, and /e/ of Middle Portuguese dropped generally, hence *bok* 'mouth' (P *boca*), *venen* /ve•ˈnen/ 'poison' (P *veneno*), sometimes taking the preceding onset with it as in *leb* 'hare' (P *lebre*), *rhos* 'face' (P *rostro*), etc. (Clements 1996: 84-85). In verbs, in contrast, the thematic vowels /a/, /e/, and /i/ remained, became final and in disyllabic items kept the stress of the etymological infinitive: cf. *katá* 'to sing' (P *cantar*), *bebé* 'to drink' (P *beber*), *irgí* 'to get up' (P *erguer*), a pattern that extended to the fourth class of *u*-final verbs borrowed from Marathi, e.g. *loṭú* 'to push' (Clements 2007 ; Clements & Luís 2014). In synchronic terms final stress in disyllabic verbs is accounted for by Clements's stress rule. In trisyllabic and longer verbs stress falls on the penultimate syllable as per the same stress rule as long as the final syllable is light (cf. *kombársa*). When inflection makes the final syllable heavy, stress returns to it: cf. *ispərme* /is•ˈpərme/ 'wring out' vs. *ispərmew* /is•pər•ˈmew/ 'wrung out' (Clements 1996: 87). It is therefore a fact that stress assignment in Korlai is blind to the POS distinction. There is nevertheless a suprasegmental difference between nominals and verbs. It is only that, owing to historical accidents, it now resides in syllabic structure — consonant-final nominals vs. vowel-final verbs (in the base form) — rather than stress placement.

4.5.3. Sri Lanka Portuguese

In Sri Lanka Portuguese, "[s]tress is non-contrastive, falling on the syllable containing the long vowel, or on the initial syllable if the word has no long vowel (e.g. *míída* 'measure' N, *midíí* 'measure' V, *mididóór* 'measurer', *jínjiviri* 'ginger')" (Smith 2013). Judging by the examples, however, there is an obvious correspondence between Portuguese stressed vowels and Sri Lanka Portuguese long vowels: compare *míída*$_{N}$ 'measure' < P *medida* /me•ˈdi•dɐ/ and *midíí*$_{V}$ 'to measure' < P *medir* /me•ˈdir/. Insofar as they come from Portuguese *r*-less infinitive forms as usual, Sri Lanka Portuguese verbs thus bear stress on their final syllables more often than do nouns. Perhaps this should lead to the conclusion that stress assignment in Sri Lanka Portuguese is POS-sensitive, with exceptions, the weight of the conclusion depending on how many exceptions there are.[30]

[30] This would make length a consequence of stress, not the other way around as Smith seems to suggest.

5. A summary of suprasegmentals and their function in PRCs

It is remarkable how well the Atlantic and part of the Asian PRCs have preserved the suprasegmental component as an indicator of POS as initiated as soon as the sixteenth century in Língua de Preto. On the Atlantic side, all languages distinguish nominals from verbs suprasegmentally, by way of stress in the Upper Guinea Creoles (Kriyol and Cape Verdean), of tone patterns in the Gulf of Guinea, of a combination of both in Papiamentu. In Asia, stress is clearly used as a POS marker in Papiá Kristang. If my conjecture about Sri Lanka Portuguese is correct, only Diu and Korlai Indo-Portuguese distinguish themselves from the other PRCs in not using suprasegmentals of prominence to discriminate POS. That they do not, however, is due to segmental changes that precluded such a use, yet did not prevent Korlai from exploiting syllabic structure to that end. To the exception of Gulf of Guinea Creoles which only accept CV syllables and therefore offer no relevant evidence, all languages also confirm that such a noun-verb distinction is related to the loss of final /r/ of infinitive forms of the lexifiers and that this loss only occurred in infinitives, which cries out for an explanation.

In search for this explanation, I now turn to an examination of the linguistic environments or substrates that may have contributed to the emergence of the Basic Variety and the subsequent creolized languages. As we shall see, the contribution is at best slight as far as the phenomena we are researching are concerned.

6. Possible influence of the linguistic environments

Such influence could a priori be detected in two dimensions: (i) the physical realization of prominence as stress (dynamic) or tone (musical); (ii) the assignment of prominence. Concerning the former, it is clear from the evidence given above that the creoles kept the dynamic realization mode of their lexifier unless pressure from a tonal environment was heavy enough to make them change it entirely or partially to musical, as was the case in the Gulf of Guinea and in Papiamentu. The real change that calls for an explanation, therefore, is in the criteria for prominence assignment.

Among the languages having POS-sensitive stress (i.e. discounting Diu and Korlai Indo-Portuguese) Guinea-Bissau-Casamance Kriyol on the one hand, Papiá Kristang and Sri Lanka Portuguese (if I am correct) on the other occupy a privileged position as far as interference is concerned, since the

possibly interfering environments — substrates and/or adstrates in traditional terms — are still in place to be observed. In the case of Guinea-Bissau-Casamance Kriyol, we are talking about Niger-Congo languages that we know were endemic to the Senegambian-Guinean area already when the Portuguese first landed there in the second half of the fifteenth century. These languages belong to two subgroups: Mande represented by one language, Mandinka, and Atlantic represented by at least twelve languages depending on what one considers distinct languages or varieties within the same language cluster (see Wilson 1989; Pozdniakov & Segerer 2016), and they are likely to have been the native languages of the first Black Africans deported to the Iberian peninsula. I will limit my survey to Balanta, Baynuk, Diola Fooñi, Manjaku, and Wolof.

Mandinka is a typical fully specified tone language including such minimal pairs as *kúlúŋò* (HHL) 'boat' vs. *kùlúŋò* (LHL) 'mortar' (Creissels & Sambou 2013: 36).[31] Stress does not seem to play any role in the language. Among the Atlantic languages, Balanta also is a typical tone language: cf. *sádá* (HH) 'alms' vs. *sàdá* (LH) 'calf' (Creissels & Biaye 2015: 30). As far as the available descriptions allow one to settle the matter, the remaining Atlantic languages make no use of tone as a suprasegmental, but only of stress, variously realized, the position of which is fixed and seems to depend on morphophonological factors. According to Sapir (1969: 9) and Doneux (1969: 10), for instance, primary stress is placed on the first syllable of stems in Diola Fooñi. The same seems to be true in Baynuk (Doneux 1969: 15), as well as in Wolof, except in disyllabic words where the first syllable is short and the second syllable is long such as *jigéen* /ɟi•geːn/ 'woman'. Then, according to Sauvageot (1965: 42), the stress pattern is perceived as being 'flat', with no significant difference between the two syllables in terms of loudness. Diallo (1981: 7) speaks about "*déplacement de l'accent*" in such cases, but he suspects it to be an "*impression*" due to the tenseness of the long vowel. As a matter of fact, Wolof is a language where stress is weakly marked (in comparison to, say, Portuguese or English), which probably correlates with the fact that it never fulfils any semantically distinctive function.

Things seem more complex in Manjaku, where tone contrasts are observed, e.g. in *màn tsɛ́p* /mãn$_{L}$-ʦəpˀ$_{H}$/ 'I left' vs. *mán tsɛ̀p* /mãn$_{H}$-ʦəpˀ$_{L}$/ 'I am leaving' (Doneux 1975: 7). Yet Doneux (1969: 19) argues that (i) tone in Manjaku is only contextual, never lexical as in Balanta and Mandinka; (ii) the

[31] The final low tone /ò/ is the determiner that is rarely absent from Mandinka nouns. As it expresses specificity rather than definiteness, it may be translated as 'the' or 'a' depending on context, which is why I omit it in my glosses.

apparent tone contrasts in the example just given can be more economically accounted for as resulting from differences in stress placement: in perfective *màn tsḗp*, the verb is stressed, whereas in the imperfective *mán tsèp*, the subject pronoun *mán* is actually a complex form cumulating the pronoun and the auxiliary *ka* which bears stress. This analysis, if correct, makes Manjaku a pitch accent language (see Hyman 1975: 230-232).

What this brief survey allows us to conclude, then, is that, given widespread plurilingualism, most if not all the people who were taken into slavery to Portugal by the end of the fifteenth century must have been familiar with languages in which stress was the only or main suprasegmental of prominence. Thanks to this, they were well equipped to perceive and memorize Portuguese stress patterns, which they preserved in the Basic Variety they evolved, *modulo* the systemic change due to infinitive /r/ dropping — which entailed no positional change, though. These patterns were then brought back to the African soil possibly by released interpreters and inherited by the Guinea-Bissau-Casamance Kriyol's founders who had no reason to alter them, being L1 or L2 speakers of the same languages as the captives had been (see Kihm & Rougé 2013).

But did the endemic languages of the Senegambian-Guinean area offer a model for the prosodic segregation of nouns and verbs? In Mandinka lexical categories are not phonologically distinguished in principle. It is a fact, however, that nouns nearly always (even in citation form) appear with a suffixed low tone determiner /-o^{L}/, which makes them end in a low or HL contour tone (see footnote 31), whereas verbs mostly end in a high tone (Creissels & Sambou 2013: 33ff.). Assuming low tone to be reinterpretable as unstressed and high tone as stressed, this might constitute a basis for the Guinea-Bissau-Casamance Kriyol distinction. One may also adduce the fact that in Balanta, Diola Fooñi and Manjaku nouns formally differ from verbs in that they begin with a member of the small set of noun class prefixes. In Wolof, on the other hand, conversion is a prevalent derivational device and noun class is not marked on the noun but on determiners, so that there is no audible difference between, e.g., *keppu* as a verb meaning 'to pinch' and *keppu* as a noun meaning 'clothes peg', both pronounced /ˈkɛ•pːu/.

Influence from the environment is even more elusive with Papiá Kristang. In Malay, the main contact language, to the extent there is stress at all, assigment is purely phonological and cannot be used to distinguish POS (Gil 2012). So much is true of the creolized varieties Singapore Malay and Ambon Malay (Aye 2013; Paauw 2013). It is certainly possible that African slaves were present in Malacca under Portuguese rule and even later, coming directly from

Africa or via Sri Lanka where a community claiming African origin, the Kaffirs, has thrived to this day (Silva & Angenot 2008), but it is unlikely they ever were numerous enough to exert a significant linguistic influence (see Cardoso 2010 for a general assessment of African presence in Portuguese India).

Finally, the language Sri Lanka Portuguese is most in contact with is Tamil (Smith 2013). Now stress in Tamil, to the extent it plays a role in the language, seems to be fixed on the first syllable of phonological words (Keane 2006). If there was any influence, it must have come from the other main language of Sri Lanka, Indo-Aryan Sinhala, which exhibits the same length-stress correlation as does Sri Lanka Portuguese, although not with the result of distinguishing nouns from verbs through stress location (Silva 2008). Be it as it may, inflection in Tamil and Sinhala is rich and distributed enough that dedicated POS markers are not required.

I now turn to languages not presently in contact with possibly substratal-adstratal environments, that is Cape Verdean, the Gulf of Guinea Creoles, and Papiamentu. As we saw, only languages belonging to the Barlavento group and Fogo in the Sotavento group show obviously POS-sensitive stress assignment, whereas only putative traces of it can be found in the Santiago variety of Sotavento. Given the late formation of the Barlavento variety during the first half of the nineteenth century (Swolkien 2013), the probability of direct contacts with continental endemic languages is nil. As for Sotavento, the most influential contact language according to Lang (2009, 2013) was Wolof, which offers no model for prosodic POS distinction as pointed out.

That Lung'ie and Santome evolved into tone languages is not surprising in contrast, as their main substrate, as we know, is Edo (Niger-Congo, Edoid), a typical two-tone language with register tones having both a lexical function to distinguish lexemes and a grammatical function to distinguish various tense-aspect distinctions (see Ladefoged 1975: 226; Hagemeijer & Ogie 2011).[32] In Edo verbs and nouns are indeed phonologically discriminated, although not tonally, but syllabically: verbs always begin with a consonant, nouns with a vowel, a remnant of former noun class prefixes, in a way similar to what we observed in a subset of Atlantic languages. This pattern was kept in Lung'ie, except that not all nouns begin with a vowel. Combining such a partial distinction with the tonal process sketched above may give us something that begins to look like an explanation.

Insofar as Angolar proceeds from an earlier stage of Santome (Maurer 2013a), Edo belongs to its linguistic environment. The most conspicuous

[32] In Ladefoged (1975), the language is called by its other name Bini.

contact language, however, is Kimbundu (Niger-Congo, Bantu) which provided numerous loanwords. Kimbundu is a tone language with two register tones, H and L. According to Xavier (2010: 83), *"*as ocorrências com o esquema tonal B [LL] são freqüentes tantos em nomes quanto nos verbos; com uma ligeira preferência nos radicais verbais*"*. As in Balanta, Manjaku, etc. nominal stems are usually preceded by a noun class prefix varying for number — i.e. morphosyntactically active, not a relic as in Edo — not present with verbal stems. (Note however this is not a foolproof touchstone, as in all noun class systems, Atlantic as well as Bantu, there is at least one class with a 'zero' exponent in the singular, so that the stem appears without a prefix.)

Edo was also effective in shaping Fa d'Ambô: among the 28 lexemes to which he attributes a Kwa origin, Granda (1985: 200-201) identifies 21 as from "Bini" (see footnote 30). Moreover, from the 30 lexemes of Bantu origin, 24 are traced back to Kikongo, also a tone language (Mufwene 2013).

According to Maurer (2013c) "Papiamentu originated in Curaçao in the second half of the 17th century in a society where the following languages or language families were most probably present: Dutch (and other northern European languages), European Portuguese and Spanish (spoken by the Sephardic Jews), Gbe, Bantu languages from the Congo-Angola area, Afro-Portuguese, and maybe also Pidgin Portuguese (spoken by soldiers and seamen at the fort)". This quasi-Babelian list includes two plausible sources for tones, Gbe and Bantu, and perhaps three if "Afro-Portuguese" already sounded like present-day Saramaccan. Yet the sociolinguistic scene, in particular the presence of already creolized varieties of Portuguese, makes it impossible clearly to separate an emerging creole from what may be considered its environment.

Summing up, we have been wondering about the possible role of contact with respect to two a priori independent phenomena: (i) the preservation of the lexifiers' stress-based suprasegmental system in the creoles or its conversion into a tone-based system (with stress possibly fulfilling some function); (ii) prosody (stress placement or tone pattern) as an emerging indicator of POS. Contact appears to be crucial with the former: only the Atlantic PRCs that were in close contact with tonal Niger-Congo languages, i.e. the Gulf of Guinea Creoles and Papiamentu, did develop tone systems.[33] Note that being in contact with two clearly tonal languages, Balanta and Mandinka, was not enough for

[33] To these one should add partially Iberian Saramaccan and, pending further investigation, Palenquero.

Kriyol to become one. The weight of the non-tonal languages in the environment prevailed.

With the second phenomenon, in contrast, no firm conclusion can be reached. In the Atlantic area no Niger-Congo contact language seems to use suprasegmentals to distinguish POS as PRCs consistently do, with the exception of Sotavento Cape Verdean. Of potential relevance may be the fact that in languages with a functional or relic noun class system prefixally expressed — thus excluding Wolof where, as already mentioned, noun class exponents attach to determiners, so that they never appear with bare nouns — the phonological form of nouns is usually distinct from that of verbs. Two circumstances weaken the force of this factor, however. One is that, as mentioned, all noun class systems include at least a 'zero' class exponent. The other is the not-to-be-excluded possibility that a verb stem will begin with a segment or a syllable that happens independently to be a noun class marker. Manjaku examples are *kəŋan* 'to be constipated', where the initial /kə/ syllable is homophonous with the class 10 exponent as in *pə$_9$-kəs* / *kə$_{10}$-kəs* 'eye(s)', or *nakəsa* 'to have a siesta' compared with *na$_1$-pene* / *ba$_2$-pene* 'healer(s)' (see Doneux 1975).

Even more detrimental to the assumption of contact influence is the presence of suprasegmental POS discrimination in the Asian area, in Papiá Kristang and possibly Sri Lanka Portuguese, where even the weak argument of frequent segmental distinction in the contact languages cannot be invoked.

Without entirely rejecting interference as a possible factor in the Atlantic area, we must therefore acknowledge that it is at best insufficient to account for the emergence of prosodic POS distinction. We must look elsewhere, that is at the lexifier, and seriously tackle the issue of final /r/-drop in infinitives, since that seems to be the crucial factor that triggered the change to POS-sensitive suprasegmentals. First, we shall turn our attention to a property of final /r/ in Portuguese and Romance languages more generally, namely its recurring instability.

7. The instability of final rhotics in Portuguese and Romance languages

Let me first restate the issue: (i) stress (or former high tone reinterpreting stress if the account for the verbal LL pattern of Gulf of Guinea creoles is correct) on the final light syllable of PRC verbs results from the dropping of Middle Portuguese infinitival final *-r*; (ii) final /r/ drop is not an across-the-board phenomenon — except of course in languages such as Lung'ie where all syllable final consonants disappeared — but it specifically targeted infinitives.

As a result verbs became prosodically distinct from nominals in a way they never are in Portuguese, moreover in an unexpected way. The major change following basification of the lexifier was indeed the quasi complete discarding of verbal synthetic inflections, which left verbal stems bare. One would therefore expect verbal stems to have become undistinguishable from nominal stems by shape alone, as in English *(a) hammer* vs. *(you) hammer*. A more moderate but still significant downsizing of synthetic inflection in French led to a similar outcome: cf. *(une) nage* '(a) swimming' vs. *(elle) nage* '(she) swims'. Even in Portuguese, by no means an inflection-challenged language, there are cases where nominal and verbal forms cannot be told apart out of context: e.g. *fala* 'speech' or 's/he speaks'. Yet the ensuing creoles insisted on creating a systematic distinction by capitalizing on the prominence contrast that remained as the sole lexical clue to the difference between, say, Guinea-Bissau-Casamance Kriyol verbal /gwar•ˈda/ 'to keep, to watch' (P *guardar*) and nominal /ˈgwar•da/ 'guardian' (P *guarda*). As we saw, the process began as soon as the Língua de Preto stage, although one cannot know to what extent it already was systematic, which explains why its outcome is found on both sides of the Atlantic as well as in Asia, assuming PRCs in both areas to rest on largely similar Basic Varieties of the lexifier.

This conjunction of an across-the-board lexical phenomenon with a narrowly targeted phonological accident raises the question of directionality of the causal chain. Was the need for a prosodic separation of verbs from nouns somehow preprogrammed in the minds of the restructurers and did it trigger the phonological pruning that made the separation possible? This would explain why only infinitive /-r/ was affected and would qualify the change as a shift-induced interference in Thomason's terminology (Thomason 2014)? Or is it the other way around: a mere accident exapted into a systematic feature?

In looking for an answer one should be reminded that final /r/ deletion is a pan-Romance phenomenon having at various periods affected the Western Romance languages as well as Romanian in the East. Catalan, for instance, has no problem with final consonants: cf. *cap* 'head', *gat* 'cat', and so forth. Yet word-final /r/ is often mute, not only in infinitive forms: cf. *cuiner* /kujˈne/ 'cook$_{masc}$', *dur* /du/ 'hard', *flor* /flɔ/ 'flower', *portar* /purˈta/ 'to carry', etc. What allows one to ascertain the presence of an underlying or floating /r/ (depending on one's favourite phonological theory) in *cuiner*, *dur* and *portar* is its appearing in such forms as *cuinera* /kujˈnerə/ 'cook$_{fem}$', *dura* /ˈdurə/ 'hard$_{fem}$' and *portar-ho* /purˈtaru/ 'to carry it', plus final stress in *cuiner* and *portar* — which would be exceptional otherwise. In *flor*, however, the final consonant seems to be purely orthographic: cf. *flors* /flɔs/ 'flowers' (see Fabra 1941; Yates

1975).[34] Occitan and French present much the same evidence as does Catalan (see Anglade 1921/1977; Zink 1990).

To the difference of what happened in Catalan, Occitan, and French, final /r/ deletion never made it into standard Portuguese, although it is a widespread feature of colloquial speech.[35] Recent sociolinguistic investigations revealed 13 to 19 per cent deletions of final /r/ in non-standard and/or dialectal varieties of European Portuguese (Mateus e Rodrigues 2004 ; Oliveira 1997 ; Oliveira 2018: 28; also see Rodrigues & Hora 2016). It is also well attested in Brazilian Portuguese. Historian of language Serafim da Silva Neto (1977 : 173) writes as follows about it : "Tais pronúncias… pertencem àquela categoria de tendências já contidas na deriva da língua que logo irrompem quando o meio social é turvo e incerto pela convivência de populações de origens diversas e a conseqüente falta de uma rígida norma lingüística".[36] Neto categorically— too categorically perhaps — rejects any possibility of interference, attributing the phenomenon exclusively to the release of a tendency no longer suppressed by a vanished or too weak norm.

Crucial, however, is the fact that Portuguese *r*-deletion is *not* POS-driven. In European Portuguese the main factor seems to be whether the following word begins with a vowel or a consonant, with only 10-15 per cent deletion in the former case versus 55-58 per cent in the latter case according to an investigation in Lisbon and Braga (Rodrigues & Hora 2016). In a study of the speech of Belo Horizonte (Brazil), Huback (2006) shows that final /r/ is seldom dropped in monosyllables such as *mar* 'sea', *dor* 'pain', etc., and that the probability of its being deleted increases with word length and frequency (also see Rennicke 2015: 47). Although the following onset also plays a role, these two factors seem sufficient to explain why "Deletion is most advanced word-finally in infinitive verb forms (e.g. *falar* 'to speak', *fazer* 'to do', *sair* 'to leave') of any number of syllables" (Rennicke 2015: 47): infinitives are rarely monosyllabic and they may well be the most frequently encountered final /r/ forms.

All these are contemporaneous observations. No evidence is given about when such phonological facts appeared in the language, and probably none can be provided owing to the lack of ancient reliable sociolinguistic inquiries. Now

[34] Worthy of note is the fact that /-r/ as an infinitive morph is categorically mute, whereas final /r/ is pronounced in a number of words such as *amor* /əˈmor/ 'love', *mar* /mar/ 'sea', etc.

[35] But it is a feature of Andalusian Spanish, a significant point since this is the variety African slaves transported to Spain were likely to be first exposed to, and we know there were exchanges between both countries in this domain.

[36] I am grateful to Joaquim Brandão de Carvalho for bringing this text to my attention.

it should be self-evident that whatever happened later than, say, the middle of the seventeenth century in European Portuguese — the only Portuguese there was at the time — cannot be of any relevance for the constitution of Língua de Preto and its legacy to the creoles.[37] The question is therefore: did word-final /r/ variably drop already in sixteenth century Portugal when Língua de Preto crystallized?

Reis (2015, Chapter 2) seems to answer the question as he claims that final /r/ deletion did affect the Portuguese of "lower" classes during the Renaissance and was not limited to African slaves. Unfortunately, he does not say on what he founds such a claim. Let us give him the benefit of the doubt and assume he is right. After all, given facts in neighbouring Romance languages, there is no reason why the socially scorned tendency to drop final /r/ might not be ancient, in which case it is likely to have affected the speech of the African slaves in Portugal, owing to their everyday contacts with lower-class Portuguese people. This, however, still does not tell us why infinitives would have turned out being especially vulnerable, since variable final /r/ deletion, if it did occur in Middle Portuguese, is not likely to have targeted POS any more than it does in contemporaneous Portuguese.

8. An accident and a selection

Three crucial factors favouring variable /r/-final deletion come out from the studies mentioned in the foregoing section: onset of the following word, length and frequency, with the latter two in turn possibly favouring infinitive forms as argued above. Now the point is that all three factors are properties of tokens, not of types. In other words, whenever in a given discourse one registers a word pronounced without the /r/ it ends with in other discourse segments, there are more chances for it to be an infinitive form than anything else. The prediction then is that many infinitives will be heard *r*-less, but by no means all, precisely because being an infinitive form does not play any role in final /r/-dropping, only ending in /r/ and being relatively long and frequently recurring in actual speech does.[38] Moreover, an irreducible individual variability has to be taken into account: all things being equal, some people fail to pronounce word-final /r/ more often than do other people, for reasons that escape any system.

[37] I am grateful to an anonymous reviewer for showing me I had to be as clear as possible in this matter in order to avoid misapprehensions of my reasoning.

[38] Critical length seems to be more than one syllable. But there probably is some amount of variation which corpus studies should bring to light.

This is not what we find in Língua de Preto and the creoles, however. Although the initial Portuguese input must have contained both *r*-less and *r*-ful forms, too many verbs obviously proceeding from former infinitives appear *r*-less irrespective of length and frequency in Língua de Preto for it to be mere variation, while in the creoles all verbs are *r*-less and stressed on the final CV syllable.[39] An accidental property has been morphologized (exapted) into a categorical marker of verbhood. Which drags us back to our initial puzzle: even assuming a chance beginning, why was word-final /r/ drop only systematized in verbs? Why was word-final /r/ kept in such words as Kriyol *pekadur* 'human being' (P *pecador* 'sinner') or Papiá Kristang *dibinyador* 'soothsayer', perhaps not so frequent, but certainly long enough? And note that according to Huback (2006) words ending in the *-dor* suffix are especially prone to losing their final /r/ — which suggests that what occurred during the basification-creolization process involved factors not involved in the phonological processes affecting Portuguese nonstandard varieties past or present.

Substratal influence being inconclusive as we saw, and the lexifier's input necessary but not sufficient, it remains for us to turn to the defining property of Língua de Preto, that is being a Basic Variety, a pidgin in other terms. I am referring to the already mentioned fact that a number of pidgin languages include morphs, the sole function of which is to distinguish nominal from verbal lexemes (Bakker 1995: 32-33). Bakker mentions the Basque pidgin used in the sixteenth-seventeenth centuries, in which nouns and adjectives end with <a>, i.e. the Basque suffixed definite article stripped from its meaning (also see Holm 1989 : 628-630; Mithun 1999: 324); the French-Icelandic Nautical Pidgin where verbs end with <er>, possibly pronounced /ər/ (also see Bakker 1989); and Russenorsk where most verbs end with <om> (/ʌm/ ?) and many nouns with <a> (also see Holm 1989 : 621-624). One may also mention Chinese Pidgin Russian, whose invariable verbs proceed from Russian imperatives, so that they almost all end with /aj/ or /i/ (Perekhvalskaya 2013); and Mediterranean Lingua Franca aka sabir, where final /r/ marks verbs, all of them from Romance infinitives (Schuchardt 1909 ; Dakhlia 2008).

Although such a device appears reasonable and would be expected to be widespread — POS distinction, especially of verbs and nouns, being a basic tenet of human language not "conceptually necessary" (see Carstairs-McCarthy 1999) — it seems to be quite rare cross-linguistically.[40] Actually, this is not

[39] Even if the input was a foreigner talk variety — for which there is no serious evidence — dropping final /r/ does not seem to have been part of it (Naro 1978; Kihm & Rougé 2013).

[40] This seems to be only true of oral language. It is a common, perhaps constant feature of sign languages in contrast (see Abner et al. 2019).

surprising. As pointed out in the introductory section, few 'full' languages, if any, are entirely devoid of inflectional or derivational morphs typically associated with particular POSs, such as Russian *-ov*, whose function, though, is not to distinguish verbs from nouns. Such a distinction is but a correlate of the meanings the morphs are associated with: only nouns may be marked for the feature set [genitive plural]. A different object is a phonological element whose only meaning is the POS of the word that includes it — as the above mentioned Esperanto endings and, more to the point, the prominence patterns of most PRCs.

Research on L2 unguided acquisition by adults — more precisely individuals past 'language puberty' — shows it to be a different process than L1 acquisition by prepubescent children still able to acquire languages perfectly almost without any formal instruction (Ellis 1994, 1997). In contrast, pubescent adolescents and adults having lost most of their learning capacities (for unclear but lamentably effective reasons) are constrained to learn a second language at an age when they are unable to do so perfectly even with explicit instructions, often not given.[41] Complex inflections in particular may prove an insurmountable hindrance.

Because of this, the approximate varieties managed by adult unguided learners — so-called interlanguages, Basic Varieties when 'fossilized' — will not include any of those POS-oriented inflections or derivations morphs which constitute a perhaps not indispensable, but certainly useful help in parsing utterances into predicates and arguments ('Who does what to whom?').[42] Also count with the fact that finding out the proper context of what sounds like a word or an expression in a situation where understanding is in itself a problem, not to say a source of anxiety or dread — never forget we are talking about enslaved people — is no easy task. Given such conditions, any signal allowing for easy distinction between prototypically predicative (verbs) and prototypically argumental constituents (nouns) will appear as a non-redundant, highly useful help with understanding utterances.

Such a signal may well have been the above-mentioned /aj/ ending the origin of which I found puzzling. Whatever its origin may be, though, it clearly functions as a marker that the item bearing it is a verb, just like the pidgin endings reviewed above do. Why is it not found in any PRC? If the above

[41] The well-known classic scenario involves decreasing UG-access (for a Portuguese-oriented review see Madeira 2016). For an alternative UG-free scenario see Wray & Grace (2007). For pidgin/creole formation and second language acquisition see Siegel (2008a, b).

[42] Or, as we saw perusing the Língua de Preto texts, they include them as frozen forms, no longer parts of structured paradigms.

sketchy description of the Portuguese sociolinguistic situation at the beginning of the sixteenth century is correct, it was not the African slaves who dropped final /r/, but they encountered a variation such that the same word uttered by a native speaker was heard now with final /r/, now without. Moreover, infinitives were especially prone to vary in this way owing to their average length (more than one syllable) and, one may assume, greater frequency in discourse than other *r*-final items.

Given this, the Africans' own contribution consisted firstly in picking out mostly infinitive forms to use as verbs, thus bypassing Portuguese too complex inflections; secondly in extending *r*-lessness to all verb occurrences, while keeping final /r/ in nominals, thereby suppressing variation in their speech.[43] It resulted in a (near-)systematic stress contrast that was in itself a sufficient signal. One should not be surprised, however, that a more obvious signal was 'experimented' for a while. Nor is it surprising that this signal, the actual frequency of which is unknown to us in any event, did not carry on in the creoles, as they developed enough synthetic and analytic morphology to separate nouns from verbs, which freed them from the need of dedicated markers for the distinction. Differential prominence, where it kept on, became the icing on the cake, if I may say so.

9. Conclusion

Like most, maybe all changes, the emergence of POS-sensitive prominence first in the original Basic Variety, then in nearly all PRCs results from the conjunction of an accident and a selection within the limits of what the overall system, here human language, allows. The accident was the variable final *r*-dropping I assume to have already been a feature of the Portuguese language in the sixteenth century, when the first African slaves set foot on the Iberian Peninsula. It was a necessary condition, albeit not a sufficient one, as shown by the fact that there is no evidence of POS-sensitive prominence in French-related creoles although *r*-drop, especially in infinitive forms, did occur in Middle and Early Modern French, and the creoles inherited from it: cf. Haitian *kouri* vs. French *courir* 'to run' variably pronounced /kuri/ in the seventeenth century. There is strong evidence, however, that already in the Middle period — and

[43] Of course, this is an idealization. The point is that one cannot be sure whether the variation appearing in Língua de Preto texts is genuine or due to the playwrights' not feeling it necessary systematically to note what was for them a stereotype. Be it as it may, variation is absent from the creoles, where the prosodic noun-verb contrast became fully grammaticalized.

possibly even earlier — French (sole among Romance languages) was no longer a word-stress language, but it had turned to the Modern French system where words are given prominence on the final syllable, not as a lexical or morphophonological property, but either for emphasis or owing to their appearing at the end of a prosodic domain (Vaissière 1996). Such a system gave the creolized offspring of French no foundation for the emergence of POS-sensitive stress patterns.

The selection, that is the conversion of a variable, contentless phenomenon into a regular grammatical feature — that verbs end in a stressed vowel, whereas nouns do not — was due to the cognitive need to ensure easy POS recognition, namely noun-verb distinction, in a context combining unguided L2 acquisition by adults with basification, that is the near total loss of those POS-associated inflections and derivations that take charge of the distinction in probably all languages and, certainly, in the native languages of the said adults. The need, I assume, is a constant across all such situations. It must have been sorely felt by people brutally transported from Africa to an entirely new world (not a brave one, to be sure).

Once a (still variable) feature of the Basic Variety, POS-sensitive prominence was inherited as a grammatical feature by most ensuing creoles, thereby giving rise to a significant typological difference between them and the lexifier. The precise routes by which this legacy travelled is something that has still to be worked out. As proposed by Kihm & Rougé (2013), interpreters may have played an important role, at least in West Africa, in disseminating the Basic Variety in the process of becoming a more stable pidgin, then creoles. Other media were probably active as well, especially in Asia. Contact languages (substrate-adstrate) do not seem to have taken a significant part in the change.

References

Abner, Natasha, Molly Flaherty, Katelyn Stangl, Marie Coppola, Diane Brentari & Susan Goldin-Meadow. 2019. The noun-verb distinction in established and emergent sign systems, *Language* 95(2), 230-267.

Agostinho, Ana Lívia. 2015. Fonologia e método pedagógico do lung'ie. Tese de Doutorado em Filologia e Língua Portuguesa. Universidade de São Paulo.

Agostinho, Ana Lívia & Larry M. Hyman. 2021. Word prosody in Lung'Ie: One system or two?, *Probus* 33(1), 57-93.

Andrade, Ernesto d' & Alain Kihm. 2000. Stress in three creoles. In E. d'Andrade, M.A. Mota & D. Pereira (org.), *Crioulos de Base Portuguesa. Actas do Workshop sobre Crioulos de Base Lexical Portuguesa (29-30 de Abril de 1999)*, 97-109. Braga: Associação Portuguesa de Linguística.

Anglade, Joseph. 1921/1977. *Grammaire de l'ancien provençal*. Paris: Klincksieck.

Aye, Khin Khin. 2013. Singapore Bazaar Malay. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, Vol. III. Oxford: Oxford University Press.

Bakker, Peter. 1989. "The Language of the Coast Tribes is Half Basque": A Basque-American Indian Pidgin in use between Europeans and Native Americans in North America ca. 1540-ca. 1640, *Anthropological Linguistics* 31(3-4), 117-147.

Bakker, Peter. 1995. Pidgins. In J. Arends, P. Muysken & N. Smith (eds), *Pidgins and Creoles: An Introduction*, 25-39. Amsterdam: John Benjamins.

Bakker, Peter. 2008. Pidgins versus creoles and pidgincreoles. In S. Kouwenberg & J.V. Singler (eds), *The Handbook of Pidgin and Creole Studies*, 130-157. Oxford: Wiley-Blackwell.

Barrena, Natalio P. 1957. *Gramática annobonesa.* Madrid: Consejo superior de Investigaciones científicas.

Baxter, Alan N. 1988. *A Grammar of Kristang (Malacca Creole Portuguese)*. Canberra : Pacific Linguistics.

Berardinelli, Cleonice & Ronaldo Menegaz (eds). 1994. *Teatro de António Ribeiro Chiado (Autos e Práticas)*. Porto: Lello & Irmão Editores.

Biagui, Noël Bernard. 2018. *Description générale du créole afro-portugais de Ziguinchor (Sénégal)*. Paris: Karthala.

Biagui, Noël Bernard & Nicolas Quint. 2013. Casamancese Creole. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, *Vol. II*, 40-49. Oxford: Oxford University Press.

Bull, Benjamim Pinto. 1989. *O Crioulo de Guiné-Bissau: Filosofia e Sabedoria*. Lisboa: Instituto de Cultura e Língua Portuguesa.

Cardoso, Eduardo Augusto. 1989. *O Crioulo da Ilha de S. Nicolau de Cabo Verde*. Praia: Ministério da Educação.

Cardoso, Hugo C. 2009a. *The Indo-Portuguese Language of Diu*. Utrecht: LOT.

Cardoso, Hugo C. 2009b. Jacques Arends' model of gradual creolization. In R. Selbach, H.C. Cardoso & M. van den Berg (eds), *Gradual Creolization: Studies celebrating Jacques Arends*, 13-23. Amsterdam: John Benjamins.

Cardoso, Hugo C. 2010. The African slave population of Portuguese India: Demographics and impact on Indo-Portuguese, *Journal of Pidgin and Creole Languages* 25(1), 95-119.

Carstairs-McCarthy, Andrew. 1999. *The Origins of Complex Language: An Inquiry into the Evolutionary Beginnings of Sentences, Syllables, and Truth*. Oxford: Oxford University Press.

Carstairs-McCarthy, Andrew. 2010. *The Evolution of Morphology*. Oxford: Oxford University Press.

Chapouto, Sandra Marisa da Costa. 2014. Contributo Para a Descrição de Aspetos Fonológicos e Prosódicos Do Crioulo Guineense. Dissertação de Mestrado em Linguística. Universidade de Coimbra.

Chasca, Edmund de. 1946. The phonology of the speech of the Negroes in early Spanish drama, *The Hispanic Review* 14(4), 322-339.

Clements, J. Clancy. 1996. *The Genesis of a Language. The Formation and Development of Korlai Portuguese*. Amsterdam : John Benjamins.

Clements, J. Clancy. 2007. Korlai (Creole Portuguese) or Nɔ Liŋ. In J. Holm & P.L. Patrick (eds), *Comparative Creole Syntax. Parallel Outlines of 18 Creole Grammars*, 153-172. London: Battlebridge Publications.

Clements, J. Clancy & Ana R. Luís. 2014. Contact intensity and the borrowing of bound morphology in Korlai Indo-Portuguese. In F. Gardani, P. Arkadiev & N. Amiridze (eds), *Borrowed Morphology*, 219-239. Berlin: De Gruyter Mouton.

Creissels, Denis & Pierre Sambou. 2013. *Le mandinka. Phonologie, grammaire, textes*. Paris: Karthala.

Creissels, Denis & Séckou Biaye. 2015. Le balant ganja. Phonologie, morphosyntaxe, liste lexicale, texte. Ms.

Dakhlia, Jocelyne. 2008. *Lingua franca : histoire d'une langue métisse en Méditerranée*. Arles : Actes Sud.

Delgado, Carlos Alberto. 2008. *Crioulo de Cabo Verde: Situação linguística da zona do Barlavento*. Praia: Instituto da Biblioteca Nacional e do Livro.

Diallo, Amadou. 1981. *Une phonologie du wolof*. Dakar : Centre de Linguistique appliquée de Dakar.

Doneux, Jean-Léonce. 1969. *Les systèmes phonologiques des langues de Casamance*. Dakar : Centre de Linguistique appliquée de Dakar.

Doneux, Jean-Léonce. 1975. *Lexique manjaku*. Dakar : Centre de Linguistique appliquée de Dakar.

Doneux, Jean-Léonce. 1979. *Systèmes phonologiques des langues de Casamance et français parlé par les élèves de CM1 (cours moyen première année)*. Dakar : Centre de Linguistique appliquée de Dakar.

Ellis, Rod. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Ellis, Rod. 1997. *Second Language Acquisition*. Oxford: Oxford University Press.

Fabra, Pompeu. 1941. *Grammaire catalane*. Paris: Les Belles Lettres.

Ferraz, Luiz I. 1979. *The Creole of São Tomé*. Johannesburg: Witwatersrand University Press.

Ferraz, Luiz I. 1984. The substrate of Annobonese, *African Studies* 43(2), 119-136.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum & Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell.

Giese, Wilhelm. 1932. Notas sóbre a fala dos negros em Lisboa no princípio do século XVI, *Revista Lusitana* XXX, 251-257.

Gil, David. 2012. A typology of stress, and where Malay/Indonesian fits in (abstract). http://email.eva.mpg.de/~gil/ismil/11/abstracts/Gil.pdf

Granda, Germán de. 1985. *Estudios de Lingüística afro-románica*. Valladolid: Universidad de Valladolid.

Granda, Germán de. 1990. Retenciones africanas en la fonética del criollo portugues de Annobón, *Papia* 1(1), 26-37.

Hagemeijer, Tjerk & Ota Ogie. 2011. Ẹ̀dó influence on Santome: Evidence from verb serialisation. In C. Lefebvre (ed.), *Creoles, their Substrates, and Language Typology*, 37-60. Amsterdam: John Benjamins.

Hagemeijer, Tjerk, Philippe Maurer-Cecchini & Armando Zamora Segorbe. 2020. *A Grammar of Fa d'Ambô*. Berlin: De Gruyter Mouton.

Holm, John. 1989. *Pidgins and Creoles. Volume II: Reference Survey*. Cambridge: Cambridge University Press.

Huback, A. P. S. 2006. Cancelamento do (r) final em nominais: uma abordagem difusionista, *Scripta* 9(18), 11–28.

Hyman, Larry M. 1975. *Phonology. Theory and Analysis*. New York: Holt, Rinehart and Winston.

Jacobs, Bart. 2013. Reassessment of the Portuguese contribution to the Papiamentu lexicon, *Journal of Pidgin and Creole Languages* 28(1), 154-165.

Jespersen, Otto. 1954. *A Modern English Grammar on Historical Principles. Part 1: Sounds and Spelling*. London and New York: Routledge.

Kalocsay, Kálmán & Gaston Waringhien. 1985. *Plena Analiza Gramatiko de Esperanto (5a korektita eldono)*. Rotterdam : Universala Esperanto-Asocio.

Keane, Elinore. 2006. Prominence in Tamil, *Journal of the International Phonetic Association* 36(1), 1-20.

Kihm, Alain. 1994. *Kriyol Syntax : The Portuguese-based Creole language of Guinea-Bissau*. Amsterdam : John Benjamins.

Kihm, Alain & Jean-Louis Rougé. 2013. *Língua de Preto*, the Basic Variety at the root of West African Portuguese Creoles: A contribution to the theory of pidgin/creole formation as second-language acquisition, *Journal of Pidgin and Creole Languages* 28(2), 203-298.

Klein, Wolfgang & Clive Perdue. 1997. The Basic Variety, or couldn't natural languages be much simpler? *Second Language Acquisition Research* 13(4), 301-347.

Kouwenberg, Silvia. 2004. The grammatical function of Papiamentu tone, *Journal of Portuguese Linguistics* 3, 55-69.

Ladefoged, Peter. 1975. *A Course in Phonetics*. New York: Harcourt Brace Jovanovich, Inc.

Lang, Jürgen. 2013. Santiago. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, *Vol. II*. Oxford: Oxford University Press.

Luís, Ana R. 2008. Tense marking and inflectional morphology in Indo-Portuguese creoles. In S. Michaelis (ed.), *Roots of Creole Structures: Weighing the contributions of substrates and superstrates*, 83-121. Amsterdam: John Benjamins.

Madeira, Ana. 2016. Second language acquisition. In W. Leo Wetzels, João Costa & Sergio Menuzzi (eds), *The Handbook of Portuguese Linguistics*, 578-590. Oxford: Wiley-Blackwell.

Magalhães, José. 2016. Main stress and secondary stress in Brazilian and European Portuguese. In W. Leo Wetzels, João Costa & Sergio Menuzzi (eds), *The Handbook of Portuguese Linguistics*, 107-124. Oxford: Wiley-Blackwell.

Mateus, Maria Helena Mira & Ernesto d'Andrade. 2000. *The Phonology of Portuguese*. Oxford: Oxford University Press.

Mateus, Maria Helena Mira & C. Rodrigues. 2004. A vibrante em coda no Português Europeu. *Actas do XIX^o Encontro Nacional da Associação Portuguesa de Linguística*, 289-299.

Maurer, Philippe. 2008. A first step towards the analysis of tone in Santomense. In S.M. Michaelis (ed.), *Roots of Creole Structures: Weighing the contribution of substrates and* superstrates, 253-261. Amsterdam: Benjamins.

Maurer, Philippe. 2009. *Principense. Grammar, Texts, and Vocabulary of the Afro-Portuguese Creole of the Island of Príncipe, Gulf of Guinea*. London: Battlebridge.

Maurer, Philippe. 2013a. Angolar. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, *Vol. II*, 59-71. Oxford: Oxford University Press.

Maurer, Philippe. 2013b. Principense. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, *Vol. II*. Oxford: Oxford University Press.

Maurer, Philippe. 2013c. Papiamentu. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, *Vol. II*. Oxford: Oxford University Press.

Meintel, Deirdre. 1975. The creole dialect of the Island of Brava. In M.F. Valkhoff (ed.), *Miscelânia Luso-Africana*, 205-256. Lisboa: Junta de Investigações Científicas do Ultramar.

Mithun, Marianne. 1999. *The Languages of Native North America*. Cambridge: Cambridge University Press.

Mufwene, Salikoko S. 2013. Kikongo-Kituba. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, *Vol. III*, 3-12. Oxford: Oxford University Press.

Naro, Anthony J. 1978. A study on the origins of pidginization. *Language* 54: 314-347.

Neto, Serafim da Silva. 1979. *História da língua portuguesa*. Rio-de-Janeiro: Presença.

Oliveira, M. A. 1997. Reanalisando o processo de cancelamento do (r) em final de sílaba, *Revista de Estudos Linguísticos* 6(2), 31–58.

Oliveira, Aline de Jesus Farias. 2018. O apagamento do rótico na (re)organização silábica. Dissertação de Mestrado, Universidade federal do Rio de Janeiro.

Paauw, Scott. 2013. Ambon Malay. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, Vol. III. Oxford: Oxford University Press.

Perekhvalskaya, Elena. 2013. Chinese Pidgin Russian. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages, Vol. III*. Oxford: Oxford University Press.

Post, Marieke. 2013. Fa d'Ambô. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, *Vol. II*, 81-89. Oxford: Oxford University Press.

Pozdniakov, Konstantin & Guillaume Segerer. 2016. Genealogical classification of Atlantic languages. In F. Lüpke (ed.), *Oxford Guide to the World's Languages : Atlantic*. Oxford : Oxford University Press.

Reis, Leonardo. 2015. *Sociolinguistic Aspects of Brazilian Portuguese : R Deletion*. Bloomington : AuthorHouse.

Remijsen, Bert & Vincent J. van Heuven. 2005. Stress, tone and discourse prominence in the Curaçao dialect of Papiamentu, *Phonology* 22, 205-235.

Rennicke, Iiris. 2015. Variation and change in the rhotics of Brazilian Portuguese. Doctoral Dissertation, University of Helsinki.

Rodrigues, Celeste & Dermerval da Hora. 2016. Main current processes of phonological variation. In W. Leo Wetzels, João Costa & Sergio Menuzzi (eds), *The Handbook of Portuguese Linguistics*, 504-525. Oxford: Wiley-Blackwell.

Rougé, Jean-Louis. 1988. *Petit dictionnaire étymologique du kriol de Guinée-Bissau et Casamance*. Bissau: Instituto Nacional de Estudos e Pesquisas.

Rougé, Jean-Louis. 1994. A propos de la formation des créoles du Cap Vert et de Guinée, *Papia : Revista de crioulos de base ibérica – Revista de criollos de base ibérica* 3(2), 137-149.

Rougé, Jean-Louis. 2004. *Dictionnaire étymologique des créoles portugais d'Afrique*. Paris : Karthala.

Rougé, Jean-Louis & Alain Kihm. 2018. Ecritures spontanées et grammaire du kriol. Paper presented at the conference of the *Associação dos Crioulos de Base Lexical Portuguesa e Espanhola* (ACBLPE), Université Assane Seck, Ziguinchor (Sénégal), 11-14 juin 2018.

Santos Morillo, Antonio. 2020. *¿Quién te lo vezó a decir? El habla de negro en la literatura del XVI, imitación de una realidad lingüística*. Berlin: Iberoamericana Vervuert.

Sapir, J. David. 1969. *A Grammar of Diola-Fogny. A language spoken in the Basse-Casamance region of Senegal*. Cambridge : Cambridge University Press.

Sauvageot, Serge. 1965. *Description synchronique d'un dialecte wolof: le parler du Dyolof*. Dakar: IFAN.

Schang, Emmanuel. 2003. L'émergence des créoles portugais du Golfe de Guinée. Thèse de doctorat, Université Nancy 2.

Schuchardt, Hugo. 1909. Die Lingua Franca, *Zeitschrift für romanische Philologie* 33, 441-461.

Siegel, Jeff. 2008a. Pidgins/Creoles, and second language acquisition. Dans Kouwenberg, S. & J.V. Singler (eds), *The Handbook of Pidgin and Creole Studies*, 189-218. Oxford: Wiley-Blackwell.,.

Siegel, Jeff. 2008b. *The Emergence of Pidgin and Creole Languages*. Oxford: Oxford University Press.

Sippola, Eeva. 2013a. Ternate Chabacano. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, *Vol. II*, 143-148. Oxford: Oxford University Press.

Sippola, Eeva. 2013b. Cavite Chabacano. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, *Vol. II*, 149-155. Oxford: Oxford University Press.

Silva, A.W.L. 2008. *Teach Yourself Sinhalese*. Colombo: Vijitha Yapa Publications.

Silva, Shihan Jayasuriya de & Jean-Pierre Angenot (eds). 2008. *Uncovering the History of Africans in Asia*. Leiden : Brill.

Smith, Ian R. 2013. Sri Lanka Portuguese. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, *Vol. II*, 111-121. Oxford: Oxford University Press.

Swolkien, Dominika. 2013. Cape Verdean Creole of São Vicente. In S.M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds), *The Survey of Pidgin and Creole Languages*, Vol. II, 20-30. Oxford: Oxford University Press.

Tavares Moreira, Ana Karina. 2020. Documentation et description grammaticale et lexicale du créole afro-portugais de l'île de Fogo (République du Cap-Vert, Afrique de l'Ouest) (Documentação e descrição gramatical e lexical do crioulo afro-português da ilha do Fogo (República de Cabo Verde, África Ocidental)). Thèse de doctorat, Institut National des Langues et Civilisations Orientales (INALCO), Paris.

Teyssier, Paul. 1959. *La langue de Gil Vicente*. Paris : Klincksieck. [2005. *A Língua de Gil Vicente*. Lisboa : Imprensa Nacional.]

Teyssier, Paul. 1980. *Histoire de la langue portugaise*. Paris : Presses Universitaires de France. [1982. *História da Língua Portuguesa* (transl. Celso Cunha). Lisboa : Martins Fontes.]

Thomason, Sarah G. 2014. When is the diffusion of inflectional morphology not dispreferred? In F. Gardani, P. Arkadiev & N. Amiridze (eds), *Borrowed Morphology*, 27-46. Berlin: De Gruyter Mouton.

Tinhorão, José Ramos. 1988. *Os Negros em Portugal: uma presença silenciosa*. Lisboa : Caminho.

Traill, Anthony & Luiz Ferraz. 1981. The interpretation of tone in Principense Creole, *Studies in African Linguistics* 12(2), 205-215.

Vaissière, Jacqueline. 1996. From Latin to Modern French: on diachronic changes and synchronic variations, *Arbeitsberichte, Institut für Phonetik und digitale Sprachverarbeitung*, Universität Kiel, 61-74.

Veiga, Manuel. 1995. *O Crioulo de Cabo Verde : Introdução à Gramática*. Praia : Instituto Nacional da Cultura.

Wetzels, W. Leo. 2006. Primary word stress in Brazilian Portuguese and the weight parameter, *Journal of Portuguese Linguistics* 5/6, 59-90.

Wetzels, W. Leo & B. Hermans. 2014. Branchingness and Stress in Korlai Portuguese. In R. Kager, J. Grijzenhout, & K. Sebregts (eds), *Where the Principles Fail. A Festschrift for Wim Zonneveld on the Occasion of his 64th Birthday,* 77-88. Utrecht : UIL/OTS.

Wilson, W.A.A. 1962. *The Crioulo of Guiné*. Johannesburg: Witwatersrand University Press.

Wilson, W.A.A. 1989. Atlantic. In J. Bendor-Samuel (ed.), *The Niger-Congo Languages*, 81-104. Lanham : University Press of America.

Wray, Alison & George W. Grace. 2007. The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form, *Lingua* 117, 543-578.

Xanthos, Aris[1], Sabine Laaha[2], Steven Gillis[3], Ursula Stephany[4], Ayhan Aksu-Koç[5], Anastasia Christofidou[6], Natalia Gagarina[7], Gordana Hrzica[8], F. Nihan Ketrez[9], Marianne Kilani-Schoch[1], Katharina Korecky-Kröll[2], Melita Kovačević[8], Klaus Laalo[10], Marijan Palmović[8], Barbara Pfeiler[11], Maria D. Voejkova[12], & Wolfgang U. Dressler[2]. 2011. On the role of morphological richness in the early development of noun and verb inflection, *First Language* 31(4) : 461-479.

Xavier, Francisco da Silva. 2010. Fonologia segmental e supra-segmental do Quimbundo. Variedades de Luanda, Bengo, Quanza Norte e Malange. Tese de Doutorado. São Paulo: Universidade de São Paulo.

Yates, Alan. 1975. *Catalan*. London: Hodder & Stoughton.

Zamora Segorbe, Armando. 2010. *Gramática descriptiva del fá d'ambô*. Malabo: Centros culturales españoles de Guinea Ecuatorial.

Zink, Gaston. 1990. *Le moyen français*. Paris: Presses Universitaires de France.

Appendix

Verb forms and tokens in three Gil Vicente's plays

1. abré	2. bebee ~ bevê (2)
3. buscá (2)	4. casaa (2)
5. chamá (2)	6. comê (2)
7. conhecê	8. dexá
9. deytá	10. dormi
11. falá (4)	12. fartá
13. fazê (2)	14. frataa
15. fruriá (2)	16. furtá ~ furutá ~ furutaa (7)
17. furugá	18. lembraa
19. metê	20. morê
21. pagá (2)	22. pari (2)
23. podê	24. puruguntaa
25. queree ~ querê (3)	26. rogá
27. trazee	

I. Verb forms marked for final stress by means of accent or vowel doubling

Forms = 27
Tokens = 47

1. buscae ~ busucay (6)	2. cansae
3. compray ~ comprae (2)	4. crivaninhae (2)
5. day	6. estae
7. falae ~ falay (4)	8. faratae
9. frutae (2)	10. furtay (3)
11. guardai ~ guarday (2)	12. mandae ~ manday (2)
13. matay	14. oyae ~ oyay (2)
15. sapantay	16. tomae ~ tomay (3)
17. tornae (2)	

II. Verb forms marked for final stress by means of apparent diphthongization

Forms = 17
Tokens = 35

1. chama (2)	2. conhece
3. contenta	4. cuyda
5. dae	6. dexa
7. dize ~dise (2)	8. fala
9. faze (2)	10. fica
11. guarda	12. leva (2)
13. manda (2)	14. promete
15. sabe	16. tira
17. toma	18. vale
19. vee (2)	20.

III. Verb forms not marked for final stress

Forms = 19
Tokens = 25

1. acendere	2. andaro ~ andar (2)
3. buscaro	4. chamar
5. chovere (2)	6. comere
7. durmir	8. fazer
9. ficara	10. levare
11. louvar	12. pedir
13. poder	14. ver

IV. Verb forms ending in -r(V)

Forms = 14
Tokens = 16

15. abre (3)	16. achasa
17. bae ~ vae ~ vay (16)	18. bem ~ vem (3)
19. cal	20. das
21. deu	22. diraa (2)
23. disse ~ rise ~ risse ~ rize (9)	24. doromia
25. durasse	26. era
27. eró ~ ro (3)	28. faraa
29. foy	30. foram
31. forcas	32. frutasse
33. furtase	34. ha (2)
35. he ~ hé (15)	36. jazia
37. matasse	38. matoo
39. queria	40. quero
41. rirá	42. ro
43. rogo (3)	44. saa ~ sae ~ sá (22)
45. saba (2)	46. sabe
47. seria	48. sey (2)
49. soo	50. tem (7)
51. temo	52. tenha
53. trae	54. vaas ~ vas (2)
55. vamo	56. vejamos
57. vejo	58.

V. Verb forms showing Portuguese finite inflections

Forms = 43
Tokens = 119

Total tokens 242
Total forms 120